美国数学会经典影印系列

AMS
AMERICAN MATHEMATICAL SOCIETY

Complex Proofs of Real Theorems

实定理的复证明

Peter D. Lax, Lawrence Zalcman

高等教育出版社

美国数学会经典影印系列
AMS
American Mathematical Society

Complex Proofs of Real Theorems

实定理的复证明

Peter D. Lax, Lawrence Zalcman

高等教育出版社·北京

图字：01-2016-2496 号

实定理的复证明
Shidingli de Fuzhengming

图书在版编目 (CIP) 数据

实定理的复证明 = Complex Proofs of Real Theorems : 英文 / (美) 彼得 · D · 拉克斯 (Peter D. Lax) , (美) 劳伦斯 · 查克曼 (Lawrence Zalcman) 著 . — 影印本 . — 北京 : 高等教育出版社 , 2017.4
ISBN 978-7-04-047000-0
Ⅰ. ①实… Ⅱ. ①彼… ②劳… Ⅲ. ①定理证明—英文 Ⅳ. ①O141
中国版本图书馆 CIP 数据核字 (2016) 第 326755 号

策划编辑 李 鹏　　责任编辑 李 鹏
封面设计 张申申　　责任印制 毛斯璐

出版发行 高等教育出版社
社址 北京市西城区德外大街4号
邮政编码 100120
购书热线 010-58581118
咨询电话 400-810-0598
网址 http://www.hep.edu.cn
http://www.hep.com.cn
网上订购 http://www.hepmall.com.cn
http://www.hepmall.com
http://www.hepmall.cn
印刷:北京新华印刷有限公司

开本 787mm × 1092mm 1/16
印张 6.75
字数 170 千字
版次 2017 年 4 月第 1 版
印次 2017 年 4 月第 1 次印刷
定价 55.00元

[物 料 号 47000-00]

美国数学会经典影印系列

出版者的话

近年来，我国的科学技术取得了长足进步，特别是在数学等自然科学基础领域不断涌现出一流的研究成果。与此同时，国内的科研队伍与国外的交流合作也越来越密切，越来越多的科研工作者可以熟练地阅读英文文献，并在国际顶级期刊发表英文学术文章，在国外出版社出版英文学术著作。

然而，在国内阅读海外原版英文图书仍不是非常便捷。一方面，这些原版图书主要集中在科技、教育比较发达的大中城市的大型综合图书馆以及科研院所的资料室中，普通读者借阅不甚容易；另一方面，原版书价格昂贵，动辄上百美元，购买也很不方便。这极大地限制了科技工作者对于国外先进科学技术知识的获取，间接阻碍了我国科技的发展。

高等教育出版社本着植根教育、弘扬学术的宗旨服务我国广大科技和教育工作者，同美国数学会（American Mathematical Society）合作，在征求海内外众多专家学者意见的基础上，精选该学会近年出版的数十种专业著作，组织出版了“美国数学会经典影印系列”丛书。美国数学会创建于1888年，是国际上极具影响力的专业学术组织，目前拥有近30000会员和580余个机构成员，出版图书3500多种，冯·诺依曼、莱夫谢茨、陶哲轩等世界级数学大家都是其作者。本影印系列涵盖了代数、几何、分析、方程、拓扑、概率、动力系统等所有主要数学分支以及新近发展的数学主题。

我们希望这套书的出版，能够对国内的科研工作者、教育工作者以及青年学生起到重要的学术引领作用，也希望今后能有更多的海外优秀英文著作被介绍到中国。

高等教育出版社

2016年12月

To our wives,
Lori and Adrienne

Preface

At the middle of the twentieth century, the theory of analytic functions of a complex variable occupied an honored, even privileged, position within the canon of core mathematics. This "particularly rich and harmonious theory," averred Hermann Weyl, "is the showpiece of classical nineteenth century analysis."[1] Lest this be mistaken for a gentle hint that the subject was getting old-fashioned, we should recall Weyl's characterization just a few years earlier of Nevanlinna's theory of value distribution for meromorphic functions as "one of the few great mathematical events in our century."[2] Leading researchers in areas far removed from function theory seemingly vied with one another in affirming the "permanent value"[3] of the theory. Thus, Clifford Truesdell declared that "conformal maps and analytic functions will stay current in our culture as long as it lasts";[4] and Eugene Wigner, referring to "the many beautiful theorems in the theory ... of power series and of analytic functions in general," described them as the "most beautiful accomplishments of [the mathematician's] genius."[5] Little wonder, then, that complex function theory was a mainstay of the graduate curriculum, a necessary and integral part of the common culture of all mathematicians.

Much has changed in the past half century, not all of it for the better. From its central position in the curriculum, complex analysis has been pushed to the margins. It is now entirely possible at some institutions to obtain a Ph.D. in mathematics without being exposed to the basic facts of function theory, and (incredible as it may seem) even students specializing in analysis often fulfill degree requirements by taking only a single semester of complex analysis. This, despite the fact that complex variables offers the analyst such indispensable tools as power series, analytic continuation, and the Cauchy integral. Moreover, many important results in real analysis use complex variables in their proofs. Indeed, as Painlevé wrote already at the end of the nineteenth century, "Between two truths of the real domain, the easiest and shortest path quite often passes through the complex

[1]Hermann Weyl, *A half-century of mathematics*, Amer. Math. Monthly **58** (1951), 523-553, p. 526.

[2]Hermann Weyl, *Meromorphic Functions and Analytic Curves*, Princeton University Press, 1943, p. 8.

[3]G. Kreisel, *On the kind of data needed for a theory of proofs*, Logic Colloquium **76**, North Holland, 1977, pp. 111-128, p. 118.

[4]C. Truesdell, *Six Lectures on Modern Natural Philosophy*, Springer-Verlag, 1966, p. 107.

[5]Eugene P. Wigner, *The unreasonable effectiveness of mathematics in the natural sciences*, Comm. Pure Appl. Math. **13** (1960), 1-14, p. 3.

domain,"[6] a claim endorsed and popularized by Hadamard.[7] Our aim in this little book is to illustrate this thesis by bringing together in one volume a variety of mathematical results whose formulations lie outside complex analysis but whose proofs employ the theory of analytic functions. The most famous such example is, of course, the Prime Number Theorem; but, as we show, there are many other examples as well, some of them basic results.

For whom, then, is this book intended? First of all, for everyone who loves analysis and enjoys reading pretty proofs. The technical level is relatively modest. We assume familiarity with basic functional analysis and some elementary facts about the Fourier transform, as presented, for instance, in the first author's *Functional Analysis* (Wiley-Interscience, 2002), referred to henceforth as [**FA**]. In those few instances where we have made use of results not generally covered in the standard first course in complex variables, we have stated them carefully and proved them in appendices. Thus the material should be accessible to graduate students. A second audience consists of instructors of complex variable courses interested in enriching their lectures with examples which use the theory to solve problems drawn from outside the field.

Here is a brief summary of the material covered in this volume. We begin with a short account of how complex variables yields quick and efficient solutions of two problems which were of great interest in the seventeenth and eighteenth centuries, viz., the evaluation of $\sum_1^\infty 1/n^2$ and related sums and the proof that every algebraic equation in a single variable (with real or even complex coefficients) is solvable in the field of complex numbers. Next, we discuss two representative applications of complex analysis to approximation theory in the real domain: weighted polynomial approximation on the line and uniform approximation on the unit interval by linear combinations of the functions $\{x^{n_k}\}$, where $n_k \to \infty$ (Müntz's Theorem). We then turn to applications of complex variables to operator theory and harmonic analysis. These chapters form the heart of the book. A first application to operator theory is Rosenblum's elegant proof of the Fuglede-Putnam Theorem. We then discuss Toeplitz operators and their inversion, Beurling's characterization of the invariant subspaces of the unilateral shift on the Hardy space H^2 and the consequent divisibility theory for the algebra $\mathcal{B}$ of bounded analytic functions on the disk or half-plane, and a celebrated problem in prediction theory (Szegő's Theorem). We also prove the Riesz-Thorin Convexity Theorem and use it to deduce the boundedness of the Hilbert transform on $L^p(\mathbb{R})$, $1 < p < \infty$. The chapter on applications to harmonic analysis begins with D.J. Newman's striking proof of Fourier uniqueness via complex variables; continues on to a discussion of a curious functional equation and questions of uniqueness (and nonuniqueness) for the Radon transform; and then turns to the Paley-Wiener Theorem, which together with the divisibility theory for $\mathcal{B}$ referred to above is exploited to provide a simple proof of the Titchmarsh Convolution Theorem. This chapter concludes with Hardy's Theorem, which quantifies the fact that a function and its Fourier transform cannot both tend to zero

[6] "Entre deux vérités du domain réel, le chemin le plus facile et le plus court passe bien souvent par le domaine complexe." Paul Painlevé, *Analyse des travaux scientifiques*, Gauthier-Villars, 1900, pp.1-2.

[7] "It has been written that the shortest and best way between two truths of the real domain often passes through the imaginary one." Jacques Hadamard, *An Essay on the Psychology of Invention in the Mathematical Field*, Princeton University Press, 1945, p. 123.

too rapidly. The final chapters are devoted to the Gleason-Kahane-Żelazko Theorem (in a unital Banach algebra, a subspace of codimension 1 which contains no invertible elements is a maximal ideal) and the Fatou-Julia-Baker Theorem (the Julia set of a rational function of degree at least 2 or a nonlinear entire function is the closure of the repelling periodic points). We end on a high note, with a proof of the Prime Number Theorem. A coda deals very briefly with two unusual applications: one to fluid dynamics (the design of shockless airfoils for partly supersonic flows), and the other to statistical mechanics (the stochastic Loewner evolution).

To a certain extent, the choice of topics is canonical; but, inevitably, it has also been influenced by our own research interests. Some of the material has been adapted from [**FA**]. Our title echoes that of a paper by the second author.[8]

Although this book has been in the planning stages for some time, the actual writing was done during the Spring and Summer of 2010, while the second author was on sabbatical from Bar-Ilan University. He thanks the Courant Institute of Mathematical Sciences of New York University for its hospitality during part of this period and acknowledges the support of Israel Science Foundation Grant 395/07.

Finally, it is a pleasure to acknowledge valuable input from a number of friends and colleagues. Charles Horowitz read the initial draft and made many useful comments. David Armitage, Walter Bergweiler, Alex Eremenko, Aimo Hinkkanen, and Tony O'Farrell all offered perceptive remarks and helpful advice on subsequent versions. Special thanks to Miriam Beller for her expert preparation of the manuscript.

Peter D. Lax
New York, NY

Lawrence Zalcman
Jerusalem, Israel

[8]Lawrence Zalcman, *Real proofs of complex theorems (and vice versa)*, Amer. Math. Monthly **81** (1974), 115-137.

Contents

Preface ix

Chapter 1. Early Triumphs 1
  1.1. The Basel Problem 1
  1.2. The Fundamental Theorem of Algebra 3

Chapter 2. Approximation 5
  2.1. Completeness of Weighted Powers 5
  2.2. The Müntz Approximation Theorem 7

Chapter 3. Operator Theory 13
  3.1. The Fuglede-Putnam Theorem 13
  3.2. Toeplitz Operators 14
  3.3. A Theorem of Beurling 22
  3.4. Prediction Theory 28
  3.5. The Riesz-Thorin Convexity Theorem 34
  3.6. The Hilbert Transform 40

Chapter 4. Harmonic Analysis 45
  4.1. Fourier Uniqueness via Complex Variables (d'après D.J. Newman) 45
  4.2. A Curious Functional Equation 46
  4.3. Uniqueness and Nonuniqueness for the Radon Transform 49
  4.4. The Paley-Wiener Theorem 54
  4.5. The Titchmarsh Convolution Theorem 57
  4.6. Hardy's Theorem 58

Chapter 5. Banach Algebras: The Gleason-Kahane-Żelazko Theorem 63

Chapter 6. Complex Dynamics: The Fatou-Julia-Baker Theorem 67

Chapter 7. The Prime Number Theorem 71

Coda: Transonic Airfoils and SLE 77

Appendix A. Liouville's Theorem in Banach Spaces 81

Appendix B. The Borel-Carathéodory Inequality 83

Appendix C. Phragmén-Lindelöf Theorems 85

Appendix D. Normal Families 87

CHAPTER 1

Early Triumphs

Nothing illustrates the extraordinary power of complex function theory better than the ease and elegance with which it yields results which challenged and often baffled the very greatest mathematicians of an earlier age. In this brief chapter, we consider two such examples: the solution of the "Basel Problem" of evaluating $\sum_1^\infty 1/n^2$ and the proof of the Fundamental Theorem of Algebra. To be sure, these achievements predate the development of the theory of analytic functions; but, even today, complex variables offers the simplest and most transparent approach to these beautiful results.

1.1. The Basel Problem

Surely one of the most spectacular applications of complex variables is the use of Cauchy's Theorem and the Residue Theorem to find closed form expressions for definite integrals and infinite sums. As an illustration, we evaluate the sums

$$\zeta(2k) = \sum_{n=1}^{\infty} \frac{1}{n^{2k}}, \qquad k = 1, 2, \dots .$$

The function

$$H(z) = \frac{2\pi i}{e^{2\pi i z} - 1}$$

is meromorphic on $\mathbb{C}$ with simple poles at the integers, each having residue 1, and no other singularities in the finite plane. It follows that if f is a function analytic near the point $z = n$ $(n \in \mathbb{Z})$, then $\operatorname{Res}(H(z)f(z), n) = f(n)$. We choose $f(z) = 1/z^{2k}$ for k fixed and consider the integral

$$I_N = \frac{1}{2\pi i} \int_{\Gamma_N} H(z) \frac{1}{z^{2k}} dz, \tag{1.1}$$

where N is a positive integer and Γ_N is the positively oriented boundary of the square with vertices at the points $(N + 1/2)(\pm 1 \pm i)$. By the Residue Theorem,

$$I_N = \sum_{n=-N}^{N} \operatorname{Res}\left(H(z)\frac{1}{z^{2k}}, n\right) = \operatorname{Res}\left(H(z)\frac{1}{z^{2k}}, 0\right) + 2\sum_{n=1}^{N} \frac{1}{n^{2k}}. \tag{1.2}$$

A routine estimate shows that H is uniformly bounded on Γ_N with bound independent of N. Thus

$$H(z)\frac{1}{z^{2k}} = O\left(\frac{1}{N^{2k}}\right) \qquad \text{on} \qquad \Gamma_N;$$

and since Γ_N has length $8N + 4$, it follows from (1.1) that

$$I_N = O\left(\frac{1}{N^{2k-1}}\right).$$

Thus $\lim\limits_{N\to\infty} I_N = 0$, so from (1.2), we obtain

$$\sum_{n=1}^{\infty} \frac{1}{n^{2k}} = -\frac{1}{2}\operatorname{Res}\left(H(z)\frac{1}{z^{2k}}, 0\right). \tag{1.3}$$

To evaluate the right hand side of (1.3) explicitly, recall that the Bernoulli numbers B_n are defined by

$$\frac{x}{e^x - 1} = \sum_{\ell=0}^{\infty} \frac{B_\ell x^\ell}{\ell!}. \tag{1.4}$$

In particular, $B_0 = 1$, $B_1 = -1/2$, $B_2 = 1/6$, $B_4 = -1/30$, $B_6 = 1/42$, $B_8 = -1/30$, $B_{10} = 5/66$, $B_{12} = -691/2730$. Now from (1.4), we have

$$H(z) = \frac{2\pi i}{e^{2\pi i z} - 1} = \sum_{\ell=0}^{\infty} \frac{B_\ell (2\pi i)^\ell z^{\ell-1}}{\ell!},$$

so that the coefficient of $1/z$ in the Laurent expansion of $H(z)/z^{2k}$ about 0 is given by

$$\operatorname{Res}\left(H(z)\frac{1}{z^{2k}}, 0\right) = \frac{(-1)^k B_{2k}(2\pi)^{2k}}{(2k)!}.$$

Plugging this into (1.3) yields

$$\sum_{n=1}^{\infty} \frac{1}{n^{2k}} = \frac{(-1)^{k+1} 2^{2k-1}}{(2k)!} B_{2k}\pi^{2k},$$

which is the desired formula. In particular, taking $k = 1$, we have

$$\zeta(2) = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

COMMENTS. 1. Evaluating the sum $\sum\limits_{n=1}^{\infty} \frac{1}{n^2}$ was a celebrated problem in the mathematics of the late seventeenth and early eighteenth centuries. Originally posed by Pietro Mengoli in 1644, it was brought to public attention by Jacob Bernoulli in his *Tractatus de Seriebus Infinitis* (1689) and became known as the Basel Problem. After many unsuccessful attempts by leading mathematicians, it was finally solved in 1735 by Leonhard Euler, who produced a rigorous proof of the result in 1741. Euler went on to discover the general formula for $\zeta(2k)$, evaluating the sums explicitly for k up to 13. Of course, Euler's arguments did not make use of complex analysis, as that subject did not yet exist.

2. Expressing $\zeta(3) = \sum\limits_{n=1}^{\infty} \frac{1}{n^3}$ in a simple closed form (or proving that no such expression exists) remains an open problem of considerable interest; ditto for higher odd powers. It is known (Apéry) that $\zeta(3)$ is an irrational number; for a proof, see [**B**].

3. An extensive array of applications of the calculus of residues are displayed in the two volumes [**MK1**], [**MK2**].

Bibliography

[B] F. Beukers, *A note on the irrationality of $\zeta(2)$ and $\zeta(3)$*, Bull. London Math. Soc. **11** (1979), 268-272.

[MK1] Dragoslav S. Mitrinović and Jovan D. Kečkić, *The Cauchy Method of Residues: Theory and Applications*, D. Reidel Publishing Co., 1984.

[MK2] Dragoslav S. Mitrinović and Jovan D. Kečkić, *The Cauchy Method of Residues: Theory and Applications*, Vol. 2, Kluwer Academic Publishers, 1993.

1.2. The Fundamental Theorem of Algebra

The Fundamental Theorem of Algebra (FTA) asserts that a nonconstant polynomial

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_0 \tag{1.5}$$

with complex coefficients must vanish somewhere in the complex plane. Eighteenth century attempts to establish this result (for polynomials with real coefficients) by such worthies as Euler, Lagrange, and Laplace all proved fatally flawed; and even the geometric proof proposed by Gauss in 1799 had a (topological) gap, which was filled only in 1920 (by Alexander Ostrowski [**O**]; cf. [**Sm**, pp. 4-5]). Thus, the first rigorous proof of the theorem, published by Argand in 1814, marks an early high water mark for nineteenth century mathematics.

Complex function theory offers a particularly efficient approach for proving FTA; and proofs using such results as Liouville's Theorem, the Maximum Principle, the Argument Principle, and Rouché's Theorem appear in the standard texts. Surprisingly, however, the simplest and shortest proof, based on the Cauchy Integral Formula for circles, does not seem to have been recorded in the textbook literature.

Proof of FTA. Let the polynomial p be given by (1.5), where $n \geq 1$ and $a_n \neq 0$. First observe that

$$\lim_{R\to\infty} |p(Re^{i\theta})| = \infty \qquad \text{uniformly in } \theta \tag{1.6}$$

since

$$|p(z)| \geq |z|^n(|a_n| - |a_{n-1}|/|z| - \cdots - |a_0|/|z|^n) > \frac{|a_n|}{2}|z|^n$$

for z sufficiently large.

Now suppose that p does not vanish on $\mathbb{C}$. Then $q = 1/p$ is analytic throughout $\mathbb{C}$ and $q(0) = 1/p(0) \neq 0$. By Cauchy's integral formula,

$$q(0) = \frac{1}{2\pi i}\int_{|z|=R} \frac{q(z)}{z}dz = \frac{1}{2\pi}\int_0^{2\pi} q(Re^{i\theta})d\theta \tag{1.7}$$

for all $R > 0$. But the integral on the right hand side of (1.7) tends to 0 by (1.6) as $R \to \infty$, and we have the desired contradiction. □

Comment. The proof given above is taken from [**Z**]; cf. [**Sc**] and the discussion in [**V**].

Bibliography

[O] Alexander Ostrowski, *Über den ersten und vierten Gaussschen Beweis des Fundamental-Satzes der Algebra*, in Carl Friedrich Gauss Werke Bd. X 2, Abh. 3, Julius Springer, 1933.

[Sc] Anton R. Schep, *A simple complex analysis and an advanced calculus proof of the fundamental theorem of algebra*, Amer. Math. Monthly **116** (2009), 67-68.

[Sm] Steve Smale, *The fundamental theorem of algebra and complexity theory*, Bull. Amer. Math. Soc. (N.S.) **4** (1981), 1-36.

[V] Daniel J. Velleman, *Editor's endnotes*, Amer. Math. Monthly **116** (2009), 857-858.

[Z] Lawrence Zalcman, *Picard's Theorem without tears*, Amer. Math. Monthly **85** (1978), 265-268.

CHAPTER 2

Approximation

Analyticity can often be exploited to advantage in the study of problems of approximation, even when the objects to be approximated are functions of a real variable. We illustrate this point in the following two sections. In each of them, an essential role is played by the following basic result from functional analysis, known as the *spanning criterion.*

SPANNING CRITERION. A point z of a normed linear space X belongs to the closed linear span Y of a subset $\{y_j\}$ of X if and only if every bounded linear functional ℓ that vanishes on the subset vanishes at z, that is,

$$(*) \qquad \ell(y_j) = 0 \qquad \text{for all} \qquad y_j$$

implies that $\ell(z) = 0$.

In particular, the linear combinations of $\{y_j\}$ span all of X if and only if no bounded linear functional ℓ satisfies (*) other than $\ell = 0$.

For the proof, based on the Hahn-Banach Theorem, see [**FA**, pp. 77-78].

2.1. Completeness of Weighted Powers

Let w be a given positive continuous function defined on $\mathbb{R}$ that decays exponentially as $|t| \to \infty$:

$$(2.1) \qquad 0 < w(t) < ae^{-c|t|}, \qquad c > 0.$$

Denote by C_0 the set of continuous functions on $\mathbb{R}$ that vanish at ∞ :

$$\lim_{|t|\to\infty} x(t) = 0.$$

Then C_0 is a Banach space under the maximum norm.

THEOREM 2.1. *The functions $t^n w(t)$, $n = 0, 1, 2, \dots$, belong to C_0; their closed linear span is all of C_0. That is, every function in C_0 can be approximated uniformly on $\mathbb{R}$ by weighted polynomials.*

PROOF. We apply the spanning criterion. Let ℓ be any bounded linear functional over C_0 that vanishes on the functions $t^n w$:

$$(2.2) \qquad \ell(t^n w) = 0, \qquad n = 0, 1, \dots .$$

Let z be a complex variable, $|\operatorname{Im} z| < c$. Then $w(t)e^{izt}$ belongs to C_0, and so

$$f(z) = \ell(we^{izt})$$

is defined in the strip $|\operatorname{Im} z| < c$. We claim that f is analytic there. For the complex difference quotients of we^{izt} tend to $iwte^{izt}$ in the norm of C_0, and so

$$f'(z) = \lim_{\delta \to 0} \frac{f(z+\delta) - f(z)}{\delta} = \lim_{\delta \to 0} \ell\left(w \frac{e^{i(z+\delta)t} - e^{izt}}{\delta}\right) = \ell(iwte^{izt}).$$

Similarly for the higher derivatives; in particular, using (2.2), we have

$$\left.\frac{d^n f}{dz^n}\right|_{z=0} = i^n \ell(wt^n) = 0, \qquad n = 0, 1, \ldots .$$

Since f is analytic, the vanishing of all its derivatives at $z = 0$ means that $f(z) \equiv 0$ in the strip; in particular,

$$f(z) = \ell(we^{izt}) = 0 \qquad \text{for all } z \text{ real.}$$

By the spanning criterion, it follows that all functions we^{izt} belong to the closed linear span of $t^n w$.

According to the Weierstrass approximation theorem, every continuous *periodic* function h is the uniform limit of trigonometric polynomials. It follows that wh belongs to the closed linear span of the functions we^{izt}, z real, hence of the functions $t^n w$. Let y be any continuous function of compact support; define x by

$$x = \frac{y}{w}. \tag{2.3}$$

Denote by h a $2p$ periodic function such that

$$x(t) \equiv h(t) \qquad \text{for } |t| < p, \tag{2.4}$$

where p is chosen so large that the support of x is contained in the interval $|t| < p$. Then

$$|x - h|_{\max} \leq |x|_{\max};$$

and so, by (2.3), (2.4), and (2.1),

$$|y - wh|_{\max} \leq ae^{-cp}|x|_{\max}.$$

This shows that as $p \to \infty$, $wh \to y$. Since wh belongs to the closed linear span of the functions $t^n w$, so does y. The functions y of compact support are dense in C_0, and the proof is complete. □

COMMENT. Let w be a nonnegative function defined on $\mathbb{R}$. The polynomials are said to be complete with respect to the weight w if for each $f \in C(\mathbb{R})$ such that

$$\lim_{|x| \to \infty} w(x)|f(x)| = 0, \tag{2.5}$$

there exists, for each $\varepsilon > 0$, a polynomial P such that

$$w(x)|f(x) - P(x)| < \varepsilon \qquad \text{for all} \quad x \in \mathbb{R}.$$

The problem of finding necessary and sufficient conditions for the polynomials to be complete with respect to w was posed by S.N. Bernstein in 1924 and solved in full generality some thirty years later by S.N. Mergelyan. Mergelyan's beautiful survey article **[M]** contains a complete account of these developments, illustrated with many illuminating examples.

To connect this with the problem considered above, observe that if the polynomials are complete with respect to the positive weight w, then every function

$g \in C_0$ can be approximated uniformly by weighted polynomials. Indeed, $f = g/w$ then satisfies (2.5), and so for each $\varepsilon > 0$, there exists a polynomial P such that

$$|g(x) - w(x)P(x)| = w(x)|f(x) - P(x)| < \varepsilon \quad \text{for all} \quad x \in \mathbb{R}.$$

Bibliography

[M] S.N. Mergelyan, *Weighted approximations by polynomials*, Amer. Math. Soc. Transl. (2) **10** (1958), 59-106.

2.2. The Müntz Approximation Theorem

According to the Weierstrass approximation theorem, any continuous function $x(t)$ on the interval $[0, 1]$ can be approximated uniformly by polynomials in t. Let n be a positive integer. Clearly, if $x(t)$ is continuous on $[0, 1]$, so is

$$y(s) = x(s^{1/n}).$$

Now $y(s)$ can be approximated arbitrarily closely in the maximum norm by polynomials $p(s)$. Setting $s = t^n$, we conclude that $x(t)$ can be approximated arbitrarily closely by linear combinations of t^{jn}, $j = 0, 1, \dots$. Thus, not all powers of t are needed in the Weierstrass approximation theorem.

Serge Bernstein posed the problem of determining those sequences of positive numbers $\{\lambda_j\}$ tending to ∞ which have the property that the closed linear span of the functions

$$\{1, t^{\lambda_1}, t^{\lambda_2}, \dots\} \tag{2.6}$$

is the space $C[0, 1]$ of all continuous functions on $[0, 1]$. After some preliminary results were obtained by Bernstein, Müntz **[M]** proved the following theorem.

THEOREM 2.2. *Let $\{\lambda_j\}$ be a sequence of distinct positive numbers tending to ∞. The functions* (2.6) *span the space $C = C[0, 1]$ if and only if*

$$\sum_{j=1}^{\infty} \frac{1}{\lambda_j} = \infty. \tag{2.7}$$

PROOF. First we show that if condition (2.7) holds, the functions in (2.6) span C. Let ℓ be a bounded linear functional on C that vanishes on all the functions (2.6):

$$\ell(t^{\lambda_j}) = 0, \qquad j = 1, 2, \dots. \tag{2.8}$$

Let z be a complex variable, $\operatorname{Re} z > 0$. For such z, the function t^z belongs to C and depends analytically on z, in the sense that

$$\lim_{\delta \to 0} \frac{t^{z+\delta} - t^z}{\delta} = (\log t)t^z$$

exists in the norm topology of C. Define

$$f(z) = \ell(t^z). \tag{2.9}$$

Then f is an analytic function of z. Furthermore, since ℓ is bounded (say $\|\ell\| \le 1$) and $|t^z| \le 1$ when $0 \le t \le 1$ and $\operatorname{Re} z > 0$, it follows from (2.9) that

$$|f(z)| \le 1 \qquad \text{for} \qquad \operatorname{Re} z > 0. \tag{2.10}$$

Relation (2.8) can be expressed as

$$f(\lambda_j) = 0. \tag{2.11}$$

Define the *Blaschke product* $B_N(z)$ by

$$B_N(z) = \prod_{j=1}^{N} \frac{z - \lambda_j}{z + \lambda_j}. \tag{2.12}$$

Then

$$B_N(\lambda_j) = 0, \qquad j = 1, 2, \dots, N; \tag{2.13a}$$

$$B_N(z) \neq 0 \qquad \text{for} \qquad z \neq \lambda_j, \quad 1 \leq j \leq N; \tag{2.13b}$$

$$|B_N(z)| \to 1 \qquad \text{as} \qquad \operatorname{Re} z \to 0; \tag{2.13c}$$

$$|B_N(z)| \to 1 \qquad \text{as} \qquad |z| \to \infty. \tag{2.13d}$$

Since the zeros of B_N are shared by f,

$$g_N(z) = \frac{f(z)}{B_N(z)} \tag{2.14}$$

is analytic in $\operatorname{Re} z > 0$. We claim that

$$|g_N(z)| \leq 1 \qquad \text{for} \qquad \operatorname{Re} z > 0. \tag{2.15}$$

Indeed, combining (2.10) and (2.13c), (2.13d), we conclude that for any $\varepsilon > 0$, $|g_N(z)| \leq 1 + \varepsilon$ for $\operatorname{Re} z = \delta$ and for $|z| = \delta^{-1}$ if δ is small enough. By the maximum principle for the analytic function g_N on the domain

$$D_\delta = \{z : |z| < \delta^{-1}, \operatorname{Re} z > \delta\},$$

we have $|g_N(z)| \leq 1 + \varepsilon$ for $z \in D_\delta$. Letting first δ and then ε tend to 0, we obtain (2.15). Let k be a positive number such that $f(k) \neq 0$; then from (2.14) and (2.15), we have

$$\left| \prod_{j=1}^{N} \frac{\lambda_j + k}{\lambda_j - k} \right| \leq \frac{1}{|f(k)|}. \tag{2.16}$$

We can write the factors on the left in (2.16) as $1 + 2k/(\lambda_j - k)$. Since $\lambda_j \to \infty$, all but a finite number of these factors are greater than 1. The uniform boundedness of the product (2.16) for all N now implies (cf. [**Ah**, p. 192]) the uniform boundedness for all N of the sum

$$\sum_{j=1}^{N} \frac{1}{\lambda_j - k}.$$

But this contradicts (2.7), so we must have $f(k) = 0$ for all $k > 0$. In view of the definition (2.9) of f and property (2.8) of ℓ, this says that any linear functional that vanishes on the functions t^{λ_j} vanishes on t^k, $k > 0$. So, by the spanning criterion, it follows that all functions t^k can be approximated uniformly on $[0, 1]$ by linear combinations of the functions $\{t^{\lambda_j}\}$. Taking, in particular, $k = 1, 2, 3, \dots$ and appealing to the Weierstrass approximation theorem, we conclude that the functions (2.6) span C.

To prove the necessity of condition (2.7), let $\{\lambda_j\}$ be a sequence of positive numbers that violates (2.7):

$$\sum_{j=1}^{\infty} \frac{1}{\lambda_j} < \infty. \tag{2.17}$$

Following Rudin [**R**, pp. 314-315], define the function

$$f(z) = \frac{z}{(2+z)^3} \prod_{j=1}^{\infty} \frac{\lambda_j - z}{2 + \lambda_j + z}. \tag{2.18}$$

Since

$$1 - \frac{\lambda_j - z}{2 + \lambda_j + z} = \frac{2 + 2z}{2 + \lambda_j + z},$$

it follows from (2.17) that the product in (2.18) converges uniformly on compact subsets of the halfplane $\operatorname{Re} z > -2$ and defines there an analytic function which vanishes only at 0 and at the points λ_j, $j = 1, 2, 3, \dots$. Moreover, $f(z)$ tends to 0 quadratically as z tends to ∞ in $\operatorname{Re} z \geq -1$:

$$|f(z)| \leq \frac{\text{const.}}{|z|^2} \qquad \text{for} \quad \operatorname{Re} z \geq -1. \tag{2.19}$$

In particular, f is absolutely integrable on the line $\operatorname{Re} z = -1$.

For $\operatorname{Re} z > -1$, we can represent $f(z)$ as a Cauchy integral

$$f(z) = \frac{1}{2\pi i} \int_{C_R} \frac{f(\zeta)}{\zeta - z} d\zeta,$$

where the positively oriented contour C_R ($R > 1 + |z|$) consists of the semicircle $\{\zeta : |\zeta + 1| = R, \operatorname{Re}\zeta \geq -1\}$ traversed from $-1 - iR$ to $-1 + iR$, followed by the interval from $-1 + iR$ to $-1 - iR$ on the line $\operatorname{Re}\zeta = -1$. Let R tend to ∞; then by (2.19), we have

$$f(z) = \frac{1}{2\pi i} \int_{-1+i\infty}^{-1-i\infty} \frac{f(\zeta)}{\zeta - z} d\zeta. \tag{2.20}$$

Now for $\operatorname{Re} w > 0$,

$$\frac{1}{w} = \int_0^1 t^{w-1} dt. \tag{2.21}$$

Taking $w = z - \zeta$ and inserting (2.21) into (2.20), we obtain

$$f(z) = \frac{1}{2\pi i} \int_{-1-i\infty}^{-1+i\infty} f(\zeta) \left(\int_0^1 t^{z-\zeta-1} dt \right) d\zeta. \tag{2.22}$$

Interchanging the order of integration (which is justified by the absolute convergence of the integrals) then yields

$$f(z) = \int_0^1 t^z \left(\frac{1}{2\pi i} \int_{-1-i\infty}^{-1+i\infty} t^{-\zeta-1} f(\zeta) d\zeta \right) dt. \tag{2.23}$$

Set $\zeta = -1 + iy$, so that the inner integral becomes

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} t^{-iy} f(-1 + iy) dy = m(t). \tag{2.24}$$

Since $|f(-1 + iy)| \leq \text{const.}/(1 + y^2)$ by (2.19), the function m defined by (2.24) is a continuous function of t on $[0, 1]$.

Inserting (2.24) into (2.23), we get

$$f(z) = \int_0^1 t^z m(t)dt,$$

which we rewrite as

$$f(z) = \ell(t^z), \tag{2.25}$$

where

$$\ell(g) = \int_0^1 g(t)m(t)dt$$

for any continuous function g on $[0,1]$. Clearly, ℓ is a bounded linear functional on $C[0,1]$. By (2.18) and (2.25),

$$\ell(t^{\lambda_j}) = f(\lambda_j) = 0$$

for each $j = 1, 2, 3, \ldots$. On the other hand, since f vanishes in the right half-plane *only* at the λ_j, the functional ℓ is not identically zero. Thus, by the spanning criterion, the functions (2.6) do not span $C[0,1]$; in fact, for $\lambda > 0$, t^λ lies in the span of $\{t^{\lambda_j}\}$ if and only if $\lambda = \lambda_j$ for some j. This completes the proof of the necessity of condition (2.7). □

COMMENT. More generally, if the λ_j (> 0) are distinct but not required to converge to ∞, a necessary and sufficient condition that the functions

$$1,\ \{t^{\lambda_j}\}, \qquad j = 1, 2, \ldots \tag{2.26}$$

span $C[0,1]$ is that

$$\sum_{j=1}^{\infty} \frac{\lambda_j}{1+\lambda_j^2} = \infty; \tag{2.27}$$

cf. [**S**], [**BE**]. For a detailed discussion of Müntz's theorem and its generalizations, see [**Al**].

For distinct complex exponents λ_j, with $\operatorname{Re}\lambda_j > 0$, Szász [**Sz**] showed that the condition

$$\sum_{j=1}^{\infty} \frac{\operatorname{Re}\lambda_j}{1+|\lambda_j|^2} = \infty, \tag{2.28}$$

which reduces to (2.27) for λ_j real, is sufficient for the functions in (2.26) to span $C[0,1]$, while

$$\sum_{j=1}^{\infty} \frac{\operatorname{Re}\lambda_j + 1}{1+|\lambda_j|^2} = \infty \tag{2.29}$$

is necessary. Now (2.28) fails to hold precisely when there exists a function bounded and analytic in the right half plane which vanishes exactly at the points $\{\lambda_j\}$ [**H**, p. 132]. In view of the arguments presented above, it is thus of considerable interest that (2.28) turns out *not* to be necessary for the functions (2.26) to span $C[0,1]$ in the case of complex exponents [**S**, pp. 165-166].

Bibliography

[Ah] Lars V. Ahlfors, *Complex Analysis*, third edition, McGraw-Hill, 1979.

[Al] J.M. Almira, *Müntz type theorems. I*, Surv. Approx. Theory **3** (2007), 152-194.

[BE] Peter Borwein and Tamás Erdelyi, *The full Müntz theorem in* $C[0,1]$ *and* $L^1[0,1]$, J. London Math. Soc. (2) **54** (1996), 102-110.

[H] Kenneth Hoffman, *Banach Spaces of Analytic Functions*, Prentice-Hall, 1962.

[M] Ch. H. Müntz, *Über den Approximationssatz von Weierstrass*, Mathem. Abhandlungen H.A. Schwarz gewidmet, Berlin, 1914, pp. 303-312.

[R] Walter Rudin, *Real and Complex Analysis*, third edition, McGraw-Hill, 1986.

[S] Alan R. Siegel, *On the Müntz-Szász theorem for* $C[0,1]$, Proc. Amer. Math. Soc. **36** (1972), 161-166.

[Sz] Otto Szász, *Über die Approximation stetiger Funktionen durch lineare Aggregate von Potenzen*, Math. Ann. **77** (1916), 482-496.

CHAPTER 3

Operator Theory

Many and various are the interactions between complex analysis and operator theory, a fact witnessed by the very existence of the autologically named journal *Complex Analysis and Operator Theory*. In this chapter, we consider a variety of applications of the theory of analytic functions to operator theory.

3.1. The Fuglede-Putnam Theorem

One particularly attractive application of complex analysis to operator theory is Marvin Rosenblum's elegant proof of the Fuglede-Putnam Theorem. Recall that a closed operator N on a complex Hilbert space is said to be normal if $N^*N = NN^*$; such an operator necessarily has a dense domain. In its full generality, the FP Theorem concerns possibly unbounded normal operators on H; but it is interesting (and nontrivial) even in the case of bounded normal operators, and that is the version we present here.

THEOREM. *Let H be a complex Hilbert space, M and N bounded normal operators on H, and B a bounded operator on H such that $BN = MB$. Then $BN^* = M^*B$.*

PROOF (Rosenblum). By induction, it follows from $BN = MB$ that $BN^k = M^kB$ for $k = 0, 1, 2, \ldots$; so for $\lambda \in \mathbb{C}$, we have

$$Be^{i\overline{\lambda}N} = B\left[\sum_{k=0}^{\infty} \frac{(i\overline{\lambda}N)^k}{k!}\right] = \left[\sum_{k=0}^{\infty} \frac{(i\overline{\lambda}M)^k}{k!}\right] B = e^{i\overline{\lambda}M}B.$$

Thus $B = e^{i\overline{\lambda}M}Be^{-i\overline{\lambda}N}$, and so

$$e^{i\lambda M^*}Be^{-i\lambda N^*} = e^{i\lambda M^*}e^{i\overline{\lambda}M}Be^{-i\overline{\lambda}N}e^{-i\lambda N^*}.$$

Since M and N are normal, this can be rewritten as

$$(*) \qquad e^{i\lambda M^*}Be^{-i\lambda N^*} = e^{i(\lambda M^* + \overline{\lambda}M)}Be^{-i(\overline{\lambda}N + \lambda N^*)}.$$

The left hand side of (*) obviously defines an entire function $F(\lambda)$ with values in the Banach algebra $B(H)$ of all bounded operators on H. The operators $\lambda M^* + \overline{\lambda}M$ and $\overline{\lambda}N + \lambda N^*$ in the exponents on the right hand side are clearly self-adjoint, so that $e^{i(\lambda M^* + \overline{\lambda}M)}$ and $e^{-i(\overline{\lambda}N + \lambda N^*)}$ are unitary and hence have norm 1. Thus F is a bounded entire function, so by Liouville's Theorem in Banach spaces (Appendix A), it is constant. But then $0 = F'(0) = i(M^*B - BN^*)$, so $M^*B = BN^*$. □

COMMENTS. 1. The proof given above also shows that if b, m, and n are elements of a C^* algebra and m and n are normal, then $bn = mb$ implies $bn^* = m^*b$.

2. For possibly unbounded normal operators M and N, the FP Theorem says that if $BN \subseteq MB$, then $BN^* \subseteq M^*B$. (Here, for operators S and T partially defined on H, $S \subseteq T$ means that for each x in the domain of S, x also belongs to the domain of T and $Sx = Tx$.) The general result follows from the bounded case via a calculation involving spectral projections of M and N. The FP Theorem was initially proved for the case $M = N$ by Bent Fuglede in 1950, answering a question posed by von Neumann in 1942, and then extended to pairs of operators the following year by C.R. Putnam. See **[R]** for these and further references.

Bibliography

[R] M. Rosenblum, *On a theorem of Fuglede and Putnam*, J. London Math. Soc. **33** (1958), 376-377.

3.2. Toeplitz Operators

We begin by recalling some basic facts from Index Theory [**FA**, pp. 300-304]. Let U and V be Banach spaces and $\mathbf{T} : U \to V$ a bounded linear map. Then $\mathbf{T}$ is said to have *finite index* if

(i) the null space $N_{\mathbf{T}}$ of $\mathbf{T}$ is a finite-dimensional subspace of U

and

(ii) the range $R_{\mathbf{T}}$ of $\mathbf{T}$ has finite codimension in V.

For such an operator, the *index* is defined as

$$\operatorname{ind} \mathbf{T} = \dim N_{\mathbf{T}} - \operatorname{codim} R_{\mathbf{T}}. \tag{3.1}$$

Two bounded linear maps $\mathbf{T} : U \to V$, $\mathbf{S} : V \to U$ are called *pseudoinverses* of one another if there exist *compact* maps $\mathbf{K} : U \to U$ and $\mathbf{H} : V \to V$ such that

$$\mathbf{ST} = \mathbf{I} + \mathbf{K} \qquad \mathbf{TS} = \mathbf{I} + \mathbf{H}.$$

A basic fact [**FA**, p. 301] is that a bounded map $\mathbf{T} : U \to V$ has finite index if and only if $\mathbf{T}$ has a pseudoinverse.

In the sequel, we shall require the following extremely useful fact [**FA**, p. 304].

Homotopy Invariance of Index. *Let $\mathbf{T}(t) : U \to V$ be a one-parameter family of bounded linear mappings, $0 \le t \le 1$. Suppose that for each t, $\mathbf{T}(t)$ has finite index and that $\mathbf{T}(t)$ depends continuously on t in the norm topology. Then* $\operatorname{ind} \mathbf{T}(t)$ *is independent of t. In particular,* $\operatorname{ind} \mathbf{T}(0) = \operatorname{ind} \mathbf{T}(1)$.

Our discussion of Toeplitz operators takes place in the Hilbert space $L^2 = L^2(S^1)$ of square integrable complex-valued functions on the unit circle S^1 with norm

$$\|u\|_2 = \left(\frac{1}{2\pi} \int_0^{2\pi} |u(\theta)|^2 d\theta \right)^{1/2}. \tag{3.2}$$

The functions $e^{ik\theta}$, $k \in \mathbb{Z}$, form an orthonormal basis: every $u \in L^2$ can be expanded as

$$u(\theta) = \sum_{k=-\infty}^{\infty} u_k e^{ik\theta}, \tag{3.3}$$

where the Fourier coefficients are given by

$$u_k = \frac{1}{2\pi}\int_0^{2\pi} u(\theta)e^{-ik\theta}d\theta, \tag{3.4}$$

and the Parseval relation

$$\|u\|_2^2 = \sum_{k=-\infty}^{\infty} |u_k|^2 \tag{3.5}$$

holds. The subspace of L^2 consisting of functions all of whose Fourier coefficients of negative index vanish is the Hardy space H^2; thus

$$u \in H^2 \quad \text{if and only if} \quad u_k = 0 \quad \text{for all} \quad k < 0. \tag{3.6}$$

The space H^2 consists of boundary values of certain functions analytic in the unit disk. Indeed,

if $f \in H^2$ has Fourier series $\sum_{n=0}^{\infty} a_n e^{in\theta}$, then $\tilde{f}(z) = \sum_{n=0}^{\infty} a_n z^n$ is an analytic function in the unit disk.

For $0 < r < 1$, the restriction of $\tilde{f}$ to the circle of radius r about 0, which we denote by $\tilde{f}_r(e^{i\theta})$, belongs to L^2; and

$$\frac{1}{2\pi}\int_0^{2\pi} |\tilde{f}_r(e^{i\theta})|^2 d\theta = \frac{1}{2\pi}\int_0^{2\pi} |\tilde{f}(re^{i\theta})|^2 d\theta = \sum_{n=0}^{\infty} |a_n|^2 r^{2n}.$$

Thus

$$\lim_{r\to 1}\frac{1}{2\pi}\int_0^{2\pi} |\tilde{f}_r(e^{i\theta})|^2 d\theta = \sum_{n=0}^{\infty} |a_n|^2 = \|f\|_2^2.$$

Moreover,

$$\|f - \tilde{f}_r\|_2^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(\theta) - \tilde{f}(re^{i\theta})|^2 d\theta = \sum_{n=0}^{\infty} |a_n|^2(1-r^n)^2,$$

so that $\lim_{r\to 1} \tilde{f}_r = f$ in L^2 norm. The orthogonal projection $\mathbf{P}_+$ of L^2 onto H^2 is defined by

$$\mathbf{P}_+ u = \sum_{k=0}^{\infty} u_k e^{ik\theta} \qquad \text{if} \qquad u(\theta) = \sum_{k=-\infty}^{\infty} u_k e^{ik\theta}. \tag{3.7}$$

Clearly, by (3.5),

$$\|\mathbf{P}_+\| = 1, \tag{3.8}$$

where $\| \; \|$ is the operator norm.

In similar fashion, we define the space H_- (also denoted by $\overline{H_0^2}$) as the subspace of L^2 consisting of functions all of whose Fourier coefficients of nonnegative index vanish and let $\mathbf{P}_-$ be the projection of L^2 onto H_-. The elements of H_- are boundary values of functions antianalytic on the unit disk, i.e., functions whose complex conjugates are analytic.

DEFINITION. Let $s(\theta)$ be a continuous complex-valued function on the unit circle S^1. We associate with s the *Toeplitz operator* $\mathbf{T}_s : H^2 \to H^2$ defined by

$$\mathbf{T}_s u = \mathbf{P}_+(su) \tag{3.9}$$

and call s the *symbol* of $\mathbf{T}_s$.

Clearly, $\mathbf{T}_s$ depends linearly on its symbol: $\mathbf{T}_{s+r} = \mathbf{T}_s + \mathbf{T}_r$.

When we represent functions of class H^2 in terms of their Fourier coefficients, a Toeplitz operator becomes a truncated discrete convolution:

$$(\mathbf{T}_s u)_k = \sum_{j=0}^{\infty} s_{k-j} u_j \qquad k = 0, 1, 2, \dots . \tag{3.9$'$}$$

Here s_n and u_n denote the nth Fourier coefficients of the functions s and u, respectively. The semi-infinite matrix in (3.9′) has identical entries along each of its dexter diagonals $k - j = \text{const}$. Such matrices are called Toeplitz matrices; they arise naturally in discretizations of partial differential operators and statistical mechanics.

Our aim is to discuss the properties of the operator $\mathbf{T}_s$, where s is a continuous complex-valued function on S^1. For such functions, we have the following result.

THEOREM 3.1. *Let s be a continuous complex-valued function on S^1 and $\mathbf{T}_s$ the Toeplitz operator with symbol s. Then $\mathbf{T}_s : H^2 \to H^2$ is a bounded operator and*

$$\|\mathbf{T}_s\| \le \max_{S^1} |s(\theta)|. \tag{3.10}$$

PROOF. Multiplication by s is obviously a bounded operator with norm bounded by the maximum value of $|s|$ on S^1; and by (3.8), $\mathbf{P}_+$ is bounded with norm 1. Since $\mathbf{T}_s$ is the composition of these two operators, we obtain (3.10). □

For symbols that do not vanish on S^1, much more can be said. To this end, recall that the winding number $W(s)$ of a curve $s(\theta)$, $0 \le \theta \le 2\pi$, about 0 can be defined geometrically as the increase in the argument of $s(\theta)$ as θ goes from 0 to 2π, divided by 2π. Since

$$\log s(\theta)\Big|_0^t = \int_0^t \frac{s'(\theta)}{s(\theta)} d\theta$$

for s continuously differentiable, for such functions this can be expressed analytically as

$$W(s) = \frac{1}{2\pi i} \int_0^{2\pi} \frac{s'(\theta)}{s(\theta)} d\theta = \frac{1}{2\pi} \operatorname{Im} \int_0^{2\pi} \frac{s'(\theta)}{s(\theta)} d\theta. \tag{3.11}$$

LEMMA 3.2. *For continuous complex-valued functions on S^1 which do not vanish,*

(i) *$W(s)$ depends continuously on s;*
(ii) *$W(s)$ takes on only integer values;*
(iii) *$W(s)$ is invariant under continuous deformation (within the class of continuous nonvanishing functions);*
(iv) *$W(s) = 0$ if and only if s has a single valued logarithm, i.e., there exists a continuous function ℓ on S^1 such that $s(\theta) = e^{\ell(\theta)}$, $0 \le \theta \le 2\pi$.*

PROOF. (i) This is obvious from the geometric definition of $W(s)$.

(ii) Since the continuously differentiable functions are uniformly dense in the continuous functions on S^1, we may, in view of (i), assume that s is continuously differentiable, so that $W(s)$ is given by (3.11). Writing

$$\varphi(t) = \int_0^t \frac{s'(\theta)}{s(\theta)} d\theta$$

for $0 \le t \le 2\pi$, we have

$$\varphi(0) = 0, \tag{3.12}$$

$$\varphi(2\pi) = 2\pi i W(s), \tag{3.13}$$

and

$$\varphi'(t) = \frac{s'(t)}{s(t)}. \tag{3.14}$$

Set

$$\Phi(t) = s(t)e^{-\varphi(t)}. \tag{3.15}$$

Then by (3.14),

$$\Phi'(t) = s'(t)e^{-\varphi(t)} + s(t)\left[-\frac{s'(t)}{s(t)}\right]e^{-\varphi(t)} = 0$$

for $0 \le t \le 2\pi$, so Φ is constant on $[0, 2\pi]$. It follows by (3.12) and (3.15) that

$$s(0) = s(0)e^{-\varphi(0)} = \Phi(0) = \Phi(2\pi) = s(2\pi)e^{-\varphi(2\pi)}$$

or

$$e^{\varphi(2\pi)} = \frac{s(2\pi)}{s(0)} = 1.$$

Hence $\varphi(2\pi)$ is an integral multiple of $2\pi i$, so by (3.13), $W(s)$ is an integer.

(iii) This follows immediately from (i) and (ii).

(iv) It suffices to prove this for s continuously differentiable since such functions are dense in the continuous functions on S^1. If $s = e^\ell$, then

$$W(s) = \int_0^{2\pi} \frac{s'(\theta)}{s(\theta)} d\theta = \int_0^{2\pi} \ell'(\theta) d\theta = \ell(2\pi) - \ell(0) = 0,$$

so s has winding number 0. On the other hand, if $W(s) = 0$, we can set

$$\ell(t) = \log s(0) + \int_0^t \frac{s'(\theta)}{s(\theta)} d\theta.$$

Clearly ℓ is continuous; and since $W(s) = 0$, $\ell(2\pi) = \ell(0)$, i.e., ℓ is a continuous function on S^1. Finally, the calculation done above for Φ shows that

$$s(t)e^{-\ell(t)} = 1 \qquad \text{for} \qquad 0 \le t \le 2\pi.$$

Thus $s = e^\ell$ on S^1, as required. □

For our discussion of the properties of $\mathbf{T}_s$, we also require the following result.

LEMMA 3.3. *For s continuous,*

$$\mathbf{C} = \mathbf{P}_+ s - s \tag{3.16}$$

is a compact map of H^2 into L^2.

PROOF. Since s is continuous, given any $\varepsilon > 0$, we can approximate s uniformly by a trigonometric polynomial s_ε so that

$$|s(\theta) - s_\varepsilon(\theta)| < \varepsilon \qquad \text{for all} \quad \theta. \tag{3.17}$$

The mapping $\mathbf{C}_\varepsilon = \mathbf{P}_+ s_\varepsilon - s_\varepsilon$ annihilates any function in H^2 of the form

$$u(\theta) = \sum_{k=M}^{\infty} u_k e^{ik\theta},$$

where M is the degree of s_ε. Since these functions form a linear subspace of H^2 of codimension M, the range of $\mathbf{C}_\varepsilon$ has dimension no greater than M. In particular, each $\mathbf{C}_\varepsilon$ is compact. It follows from (3.10) and (3.17) that $\mathbf{C}_\varepsilon$ tends to $\mathbf{C}$ uniformly in norm. Since the uniform limit of compact maps is compact, (3.16) is compact. □

We also need the following result.

LEMMA 3.4. *Within the class of continuous, complex-valued, nonvanishing functions on S^1, two functions can be continuously deformed into one another if and only if they have the same winding number.*

PROOF. The invariance of the winding number under deformation is (iii) of Lemma 3.2. To prove the opposite direction, consider first the case in which the winding number of s is zero. Such a function has a single valued logarithm $\log s(\theta)$. Deform this function to zero as $t \log s(\theta)$. Exponentiation yields

$$s(\theta, t) = e^{t \log s(\theta)} \qquad 1 \geq t \geq 0,$$

a deformation of $s(\theta)$ into the constant function 1.

Given s of winding number N, we write it as

$$s(\theta) = e^{iN\theta}(e^{-iN\theta}s(\theta)).$$

The second factor has winding number zero and therefore can be deformed into the constant function 1. So $s(\theta)$ can be deformed into $e^{iN\theta}$, $N = W(s)$. □

We can now prove the following important result.

THEOREM 3.5. *Let s be a continuous, complex-valued function which does not vanish on S^1. Then the Toeplitz operator $\mathbf{T}_s$ has finite index given by*

$$\operatorname{ind} \mathbf{T}_s = -W(s). \tag{3.18}$$

PROOF. To prove that $\mathbf{T}_s$ has finite index, it suffices to show that $\mathbf{T}_s$ has a pseudoinverse; we claim that $\mathbf{T}_{s^{-1}}$ is a pseudoinverse of $\mathbf{T}_s$. Indeed, we have

$$\mathbf{T}_{s^{-1}}\mathbf{T}_s = \mathbf{P}_+ s^{-1}\mathbf{P}_+ s = \mathbf{P}_+ s^{-1}(s + \mathbf{P}_+ s - s) = \mathbf{I} + \mathbf{P}_+ s^{-1}\mathbf{C},$$

where $\mathbf{C}$ is given by (3.16). Now $\mathbf{C}$ is compact by Lemma 3.3; thus, $\mathbf{T}_{s^{-1}}\mathbf{T}_s$ differs from the identity by a compact operator. Since s and s^{-1} play symmetric roles, it follows that $\mathbf{T}_s$ and $\mathbf{T}_{s^{-1}}$ are pseudoinverses.

To prove (3.18), let us first consider the case $s(\theta) = e^{iN\theta}$. For N positive, the Toeplitz operator $\mathbf{T}_N$ whose symbol is $e^{iN\theta}$ is just multiplication by $e^{iN\theta}$. Clearly, this has only the trivial nullspace; and its range in H^2 has codimension N, since it consists of functions of the form $\sum_{k=N}^{\infty} u_k e^{ik\theta}$. Therefore,

$$\operatorname{ind} T_N = -N. \tag{3.19}$$

For $N < 0$, the mapping $T_N = \mathbf{P}_+ e^{iN\theta}$ is onto H^2; its nullspace consists of linear combinations of $1, e^{i\theta}, \ldots$, and $e^{i(-N-1)\theta}$, and thus has dimension $-N$. Therefore, (3.19) holds for $N < 0$ as well.

We have shown in Lemma 3.4 that every nonvanishing function $s(\theta)$ of winding number N can be deformed into $e^{iN\theta}$; that is, there is a one parameter family $s(\theta, t)$, continuous in θ, t, such that

$$s(\theta, t) \neq 0, \quad s(\theta, 0) = s(\theta), \quad \text{and} \quad s(\theta, 1) = e^{iN\theta}.$$

Since the winding number $W(s)$ is invariant under continuous deformations,

$$W(s) = W(s(0)) = W(s(1)) = N. \tag{3.20}$$

It follows from (3.10) that

$$\|\mathbf{T}_{s(t)} - \mathbf{T}_{s(t')}\| = \|\mathbf{T}_{s(t)-s(t')}\| \le \max_{t,t' \in S^1} |s(t) - s(t')|.$$

Since $s(\theta, t)$ depends continuously on t, $\mathbf{T}_{s(t)}$ depends continuously on t in the norm topology. Appealing to the homotopy invariance of the index, we conclude that

$$\operatorname{ind} \mathbf{T}_s = \operatorname{ind} \mathbf{T}_N.$$

Combining this with (3.19) and (3.20), we obtain (3.18). This completes the proof. □

In the course of proving Theorem 3.5, we have shown that for the special function $s_N(\theta) = e^{iN\theta}$, the dimension of the nullspace of $\mathbf{T}_N$ is either 0 or N, depending on the sign of N. This turns out to be true for all functions s.

THEOREM 3.6. *Let s be a continuous, complex-valued, nowhere zero function on the unit circle S^1 and $\mathbf{T}_s$ the Toeplitz operator with symbol s.*

(i) *If $W(s) = 0$, then $\mathbf{T}_s$ is invertible.*
(ii) *If $W(s) > 0$, then $\mathbf{T}_s$ is one-to-one and has range of codimension $W(s)$.*
(iii) *If $W(s) < 0$, then $\mathbf{T}_s$ has a nullspace of dimension $-W(s)$ and maps H^2 onto H^2.*

PROOF. (i) As noted in the proof of Lemma 3.4, when $W(s) = 0$, s has a single-valued logarithm:

$$s(\theta) = \exp \ell(\theta) \qquad \ell(\theta) = \log s(\theta).$$

Split ℓ into its analytic and antianalytic parts:

$$\ell = \ell_+ + \ell_-, \qquad \ell_+ \in H^2, \qquad \ell_- \in H_-.$$

We assume first that s is smooth, say C^∞; then so is ℓ, and so are ℓ_+ and ℓ_-. Exponentiate to obtain

$$s = e^\ell = e^{\ell_+ + \ell_-} = e^{\ell_+} e^{\ell_-} = s_+ s_-. \tag{3.21}$$

The function s_+ is the boundary value of an analytic function and s_- the boundary value of an antianalytic function. Both are continuous up to the boundary and nonzero in the closed unit disk. We now show how to invert $\mathbf{T}_s$ with the help of s_+ and s_-. Write

$$\mathbf{T}_s u = \mathbf{P}_+ s u = f,$$

for $u, f \in H^2$. This equation means that

$$su = f + g_-, \qquad g_- \in H_-.$$

Expressing s as $s_+ s_-$ and dividing by s_-, we get

$$s_+ u = s_-^{-1} f + s_-^{-1} g_-. \tag{3.22}$$

Clearly, $s_+ u \in H^2$; moreover, since $s_-^{-1} = \exp(-\ell_-)$, the product $s_-^{-1} g_-$ belongs to H_-. Thus, applying $\mathbf{P}_+$ to (3.22) gives

$$s_+ u = \mathbf{P}_+ s_-^{-1} f,$$

so that

$$u = s_+^{-1}\mathbf{P}_+ s_-^{-1} f. \tag{3.23}$$

This shows that $s_+^{-1}\mathbf{P}_+ s_-^{-1}$ is the inverse of $\mathbf{T}_s$.

Now suppose that s is merely continuous on S^1. For any $\varepsilon > 0$, we can approximate s uniformly by a smooth function r so that

$$\max_{S^1} |s(\theta) - r(\theta)| < \varepsilon. \tag{3.24}$$

For ε sufficiently small,

$$\max_{S^1} |r^{-1}(\theta)s(\theta) - 1| < 1. \tag{3.25}$$

We draw two conclusions from this inequality.

First of all, it follows from (3.25) combined with (3.10) that

$$\|\mathbf{T}_{r^{-1}s} - \mathbf{I}\| < 1. \tag{3.26}$$

This implies that $\mathbf{T}_{r^{-1}s}$ is invertible. Indeed, writing for convenience $\mathbf{T} = \mathbf{T}_{r^{-1}s}$, we have $\mathbf{T} = \mathbf{I} - (\mathbf{I} - \mathbf{T})$. By (3.26), the series $\sum\limits_{n=0}^{\infty} (\mathbf{I} - \mathbf{T})^n$ converges in (operator) norm, and its limit is easily seen to be $\mathbf{T}^{-1}$; cf. [**FA**, p. 194].

Moreover, it also follows from (3.25) that r and s have the same winding number. Indeed, (3.25) asserts that the curve $s(\theta)/r(\theta)$ is contained in the open disk of radius 1 centered at 1, from which it is obvious that it cannot surround the origin; thus $W(r^{-1}s) = 0$. But for smooth s, we have

$$W(r^{-1}s) = \frac{1}{2\pi i}\int_0^{2\pi} \left(\frac{s}{r}\right)' \Big/ \left(\frac{s}{r}\right) d\theta = \frac{1}{2\pi i}\int_0^{2\pi} \left[\frac{s'(\theta)}{s(\theta)} - \frac{r'(\theta)}{r(\theta)}\right] d\theta = W(s) - W(r).$$

By Lemma 3.2, this persists for s merely continuous.

Therefore, since we have assumed that $W(s) = 0$, also $W(r) = 0$. Since r is smooth, it can be factored as in (3.21) $r = r_+ r_-$, where r_+ is the boundary value of an analytic function which is nowhere zero in the unit disk and r_- is the boundary value of a nowhere zero antianalytic function in the unit disk. Hence, by the argument principle, $W(r_+) = 0 = W(r_-)$.

We claim that the operator $\mathbf{T}_{r^{-1}s}$ can be factored as follows:

$$\mathbf{T}_{r^{-1}s} = \mathbf{T}_{r_-^{-1} s r_+^{-1}} = \mathbf{P}_+ r_-^{-1} s r_+^{-1} = \mathbf{P}_+ r_-^{-1} \mathbf{P}_+ s \mathbf{P}_+ r_+^{-1} = \mathbf{T}_{r_-^{-1}} \mathbf{T}_s \mathbf{T}_{r_+^{-1}}.$$

This is so because the operator $\mathbf{P}_+$ to the left of r_+^{-1} acts as the identity, while the operator $\mathbf{P}_+$ to the left of s removes an antianalytic function that would have been removed by the leftmost operator $\mathbf{P}_+$. As observed above, the operator $\mathbf{T}_{r^{-1}s}$ on the left is invertible; so are the operators $\mathbf{T}_{r_-^{-1}}$ and $\mathbf{T}_{r_+^{-1}}$ on the right because the winding numbers of r_+ and r_- are zero. It follows that the third operator in the product on the right, $\mathbf{T}_s$, is invertible too. This completes the proof of (i).

We now turn to the proof of (ii) and (iii). Denote the winding number of s by W. The function $se^{-iW\theta}$ has winding number 0; therefore, by (i), the mapping $u \to f$ given by

$$\mathbf{P}_+ s e^{-iW\theta} u = f$$

is invertible. This is the same as saying that $\mathbf{T}_s$ maps $e^{-iW\theta}H^2$ one-to-one onto H^2. From this, (ii) and (iii) follow. $\square$

COMMENTS. 1. The proof of Theorem 3.6 is due to by Gohberg, who pointed out that it also applies to piecewise continuous functions s, provided that there is some continuous function r such that inequality (3.25) is satisfied for some constant on the right less than 1.

2. More generally, the Toeplitz operator $\mathbf{T}_s$ can be defined via (3.9) for arbitrary functions $s \in L^\infty$. The extensive theory for such operators and additional generalizations are discussed in detail in [**BS**].

3. An important extension of the theory of Toeplitz operators, in which S^1 is replaced by $\mathbb{R}$, was given by Wiener and Hopf [**WH**]; cf. [**PW**, pp. 49-58] and, for the further development of that theory, [**K**]. The theories of Toeplitz operators and Wiener-Hopf operators developed in parallel until Rosenblum [**R**] noticed that the two classes of operators are unitarily equivalent. In fact, as shown subsequently by Devinatz [**D**], conformal mapping of the unit disk onto the upper half-plane establishes a unitary equivalence between a Toeplitz operator and the Fourier transform of a Wiener-Hopf operator.

4. Krein and Gohberg [**GK**] have extended Theorem 3.5 to continuous $n \times n$ *matrix-valued* functions $\mathbf{S}(\theta)$ acting by multiplication on vector-valued functions $u(\theta)$. For fixed n, denote by H^2 the subspace of L^2 vector-valued functions on S^1 whose negative Fourier coefficients are all 0. Let $\mathbf{P}_+$ be the orthogonal projection of L^2 onto H^2. Then $\mathbf{T_S} = \mathbf{P}_+\mathbf{S}$ is a bounded mapping of H^2 into H^2. Krein and Gohberg show that if $\mathbf{S}(\theta)$ is invertible at each point of S^1, then $\mathbf{T_S}$ has $\mathbf{T}_{\mathbf{S}^{-1}}$ as a pseudoinverse; the determinant $\det \mathbf{S}(\theta)$ is nonzero on S^1; and $\operatorname{ind} \mathbf{T_S} = -W(\det \mathbf{S})$. On the other hand, Theorem 3.6 is no longer true in general for matrix-valued symbols. However, when $\mathbf{S}(\theta)$ can be factored as $\mathbf{S} = \mathbf{S}_-\mathbf{S}_+$, where $\mathbf{S}_-$ is antianalytic, $\mathbf{S}_+$ analytic, and both are invertible at every point of the unit disk, one has $\mathbf{T}_\mathbf{S}^{-1} = \mathbf{S}_+^{-1}\mathbf{P}_+\mathbf{S}_-^{-1}$. Unfortunately, even when such a factorization exists, it can no longer be performed by taking logarithms. A method which yields the desired factorization for a dense open set of C^∞ matrix functions satisfying $W(\det \mathbf{S}) = 0$ by solving a Dirichlet problem for a system of nonlinear partial differential equations has been given by Lax [**L**].

Bibliography

[BS] Albert Böttcher and Bernd Silbermann (with Alexei Karlovich), *Analysis of Toeplitz Operators*, second edition, Springer-Verlag, 2006.

[D] Allen Devinatz, *On Wiener-Hopf operators*, Functional Analysis, Thompson Book Co., Washington, D.C., 1967, pp. 81-118.

[GK] I.C. Gohberg and M.G. Krein, *Systems of integral equations on a half-line with kernels depending on the difference of arguments*, Amer. Math. Soc. Transl. (2) **14** (1960), 217-287.

[K] M.G. Krein, *Integral equations on half line with kernel depending upon the difference of the arguments*, Amer. Math. Soc. Transl. (2) **22** (1962), 163-288.

[L] Peter D. Lax, *On the factorization of matrix-valued functions*, Comm. Pure Appl. Math. **29** (1976), 683-688.

[PW] Raymond E.A.C. Paley and Norbert Wiener, *Fourier Transforms in the Complex Domain*, Amer. Math. Soc., 1934.

[R] Marvin Rosenblum, *A concrete spectral theory for self-adjoint Toeplitz operators*, Amer. J. Math. **87** (1965), 709-718.

[WH] Norbert Wiener and Eberhard Hopf, *Über eine Klasse singulären Integral-gleichungen*, Sitzber. Preuss. Akad. Wiss. Berlin Phys.-Math. Kl. **30/32** (1931), 696-706.

3.3. A Theorem of Beurling

Let $\mathcal{H}$ be a separable Hilbert space with complete orthonormal basis $\{e_n\}_{n=0}^{\infty}$. Then each $x \in \mathcal{H}$ has a unique representation

$$x = \sum_{n=0}^{\infty} a_n e_n, \tag{3.27}$$

where the coefficients $a_n \in \mathbb{C}$ satisfy

$$\sum_{n=0}^{\infty} |a_n|^2 < \infty; \tag{3.28}$$

and for every sequence $\{a_n\}$ satisfying (3.28), (3.27) defines an element of $\mathcal{H}$. Consider the discrete unilateral shift of multiplicity one defined on $\mathcal{H}$, i.e., the linear operator $\mathbf{T}$ which maps e_n to e_{n+1} for each nonnegative integer n. Then $\mathbf{T}$ is clearly an isometry of $\mathcal{H}$, so that $\|\mathbf{T}\| = 1$. What are the closed invariant subspaces of $\mathbf{T}$, i.e., the closed subspaces $\mathcal{N} \subset \mathcal{H}$ such that $\mathbf{T}(\mathcal{N}) \subset \mathcal{N}$?

This question was considered, and solved, by Arne Beurling in his seminal paper [**B**]. The key to Beurling's solution is to represent $\mathcal{H}$ as a space of analytic functions on the unit disk. To this end, consider the space H of analytic functions

$$f(z) = \sum_{n=0}^{\infty} a_n z^n, \qquad |z| < 1, \tag{3.29}$$

where

$$\sum_{n=0}^{\infty} |a_n|^2 < \infty.$$

Then for $0 \le r < 1$,

$$\frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^2 d\theta = \sum_{n=0}^{\infty} |a_n|^2 r^{2n}.$$

We define the norm in H by

$$\|f\|_2^2 = \sup_{0 \le r < 1} \frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^2 d\theta = \sum_{n=0}^{\infty} |a_n|^2.$$

For $f \in H$ and $0 < r < 1$, $f_r(e^{i\theta}) = f(re^{i\theta})$ is a function in $L^2(S^1)$ by (3.29). Moreover,

$$\frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta}) - f(se^{i\theta})|^2 d\theta = \sum_{n=0}^{\infty} |a_n|^2 (r^n - s^n)^2,$$

which shows that as $r \to 1$, the functions f_r converge in $L^2(S^1)$. This limit is the *boundary value* function of $f(z)$ on the unit circle,

$$f(e^{i\theta}) = \sum_{n=0}^{\infty} a_n e^{in\theta}, \tag{3.30}$$

where, by the Riesz-Fischer Theorem, the series converges in the L^2 sense. Its L^2 norm is the norm of f in H

$$\|f\|_2^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(e^{i\theta})|^2 d\theta. \tag{3.31}$$

Thus, functions in H, defined initially on the unit disk, are in one-to-one isometric correspondence with their boundary values on the unit circle; and H is a Hilbert space with inner product

$$(f, g) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{i\theta})\overline{g(e^{i\theta})} d\theta.$$

Of course, as an element of $L^2(S^1)$, the boundary value of a function in H is defined pointwise only almost everywhere (a.e.). Accordingly, equalities involving such functions are to be understood in the L^2 sense and will, in general, hold pointwise only a.e. on S^1.

It should be clear by now that the collection of boundary functions (3.30) of functions in H is precisely the space H^2 discussed in the previous section. For convenience of notation, throughout this section, we continue to refer to this space simply as H and view its elements as functions on the disk or the circle as is convenient.

Associating to each $x \in \mathcal{H}$ given by (3.27) the corresponding function f defined by (3.29) evidently establishes an isometric isomorphism between $\mathcal{H}$ and H, under which the unilateral shift on $\mathcal{H}$ becomes the operator of multiplication by the function z on the space H.

Denote by $\mathcal{B}$ the algebra of bounded analytic functions on the open unit disk Δ with the sup norm

$$\|b\|_\infty = \sup_\Delta |b(z)|.$$

Clearly, if $b \in \mathcal{B}$,

$$\sup_{0 \leq r < 1} \left(\frac{1}{2\pi} \int_0^{2\pi} |b(re^{i\theta})|^2 d\theta \right)^{1/2} \leq \|b\|_\infty < \infty;$$

thus $\mathcal{B} \subset H$, and each $b \in \mathcal{B}$ has L^2 boundary values on the unit circle. Since $b(re^{i\theta}) \to b(e^{i\theta})$ in L^2 as $r \to 1$, we have for some sequence $r_n \to 1$, $b(r_n e^{i\theta}) \to b(e^{i\theta})$ a.e.; so the boundary value function of b is bounded (in essential sup) by $\|b\|_\infty$.

In the opposite direction, we have the following result.

THEOREM 3.7. *If the boundary values of a function $f \in H$ are essentially bounded, then f belongs to $\mathcal{B}$.*

PROOF. Assume f is given by (3.29) with boundary function $f(e^{i\theta})$ as in (3.30). Then for $0 \leq r < 1$, we have

$$f(re^{i\theta}) = \frac{1}{2\pi} \int_0^{2\pi} f(e^{it}) P_r(\theta - t) dt, \tag{3.32}$$

where the Poisson kernel, defined by

$$P_r(\theta) = \operatorname{Re} \frac{1 + re^{i\theta}}{1 - re^{i\theta}} = \frac{1 - r^2}{1 - 2r\cos\theta + r^2},$$

satisfies

$$P_r(\theta) > 0, \quad \frac{1}{2\pi} \int_0^{2\pi} P_r(\theta) d\theta = 1, \tag{3.33}$$

and, more generally,

$$\frac{1}{2\pi} \int_0^{2\pi} P_r(\theta) e^{in\theta} d\theta = r^{|n|}, \quad n = 0, \pm 1, \pm 2 \ldots. \tag{3.34}$$

Indeed, (3.32) follows immediately from (3.30) and (3.34). It is now evident from (3.32) and (3.33) that if $|f(e^{it})| \leq M$ a.e. on S^1, then $|f(z)| \leq M$ for all $z \in \Delta$. $\square$

REMARK. In general, an analytic function on Δ whose radial boundary values are essentially bounded need not belong to $\mathcal{B}$. A simple example of such a function is $f(z) = \exp[(1+z)/(1-z)]$.

For $b \in \mathcal{B}$, the operation of multiplication by b is a bounded operator on H. Indeed, writing $B(f) = bf$, we have

$$\|B\| = \sup_f \|B(f)\|_2 = \sup_f \|bf\|_2 \leq \sup_f \|b\|_\infty \|f\|_2 = \|b\|_\infty,$$

where the supremum is taken over all f such that $\|f\|_2 \leq 1$. In fact, it is not difficult to see that $\|B\| = \|b\|_\infty$.

Beurling's solution of the invariant subspace problem for the unilateral shift operator may now be stated as follows.

THEOREM 3.8. *Let N be a closed subspace of H that is invariant under multiplication by z. Then*

$$N = pH,$$

where p is a function in $\mathcal{B}$ such that

$$|p(e^{i\theta})| = 1.$$

The function p is unique up to a complex constant factor of absolute value 1.

PROOF. We claim that zN is a proper subspace of N. For otherwise, any $f \in N$ could be written

$$f = zf_1 = z^2 f_2 = \dots.$$

Viewing the functions as being defined on the disk Δ, we see this would mean that f has a zero of infinite order at the origin, an impossibility for an analytic function.

By (3.31), multiplication by z is clearly an isometry of H; therefore, zN is a proper closed subspace of N. Denote its orthogonal complement in N by M, so that

$$N = M \oplus zN. \tag{3.35}$$

Since multiplication by z preserves orthogonality, replacing N on the right by its orthogonal decomposition given by (3.35) and iterating, we obtain

$$N = M \oplus zM \oplus z^2 M \oplus \dots \oplus z^{k-1} M \oplus z^k N \tag{3.36}$$

for each k. Letting $k \to \infty$ then shows that

$$N \supset M \oplus zM \oplus z^2 M \oplus \dots.$$

We claim that the orthogonal sum on the right hand side is actually equal to N. Indeed, otherwise there would exist $g \in N$ that is orthogonal to every $z^j M$. By (3.36), such a g would belong to $z^k N$ for every k and thus would have a zero of infinite order at 0, which is impossible. Thus, in fact,

$$N = M \oplus zM \oplus z^2 M \oplus \dots. \tag{3.37}$$

Let us now examine the space M. Let $m \in M$; then by (3.36), m is orthogonal to $z^k N$, $k \geq 1$, and so, in particular, to $z^k m$

$$(z^k m, m) = \frac{1}{2\pi} \int_0^{2\pi} e^{i\theta k} |m(e^{i\theta})|^2 d\theta = 0, \quad k = 1, 2, \dots. \tag{3.38}$$

Taking complex conjugates shows that (3.38) holds for $k = -1, -2, \ldots$, as well. Thus, all Fourier coefficients of $|m(e^{i\theta})|^2$ except the zeroth vanish, which implies that $|m(e^{i\theta})|$ is constant.

We claim that M is one-dimensional. To see this, let m and p be two functions in M. Then $m + \alpha p \in M$ for any constant α; so, by what has been shown above,

$$|m + \alpha p|^2 = (m + \alpha p)(\overline{m} + \overline{\alpha}\,\overline{p}) = |m|^2 + |\alpha|^2|p|^2 + 2\,\mathrm{Re}\,\alpha p\overline{m}$$

is constant. Since α is an arbitrary complex constant, $p\overline{m}$ is constant. Dividing by $|m|^2 = m\overline{m}$, we conclude that p/m is constant, i.e., p and m are proportional.

Normalize $p(e^{i\theta})$ in M to have $|p| = 1$; then all functions in M are multiples of p. Putting this into (3.37) shows that every function $f \in N$ can be decomposed as

$$f = a_0 p + z a_1 p + \cdots = p(a_0 + a_1 z + \ldots) = pg. \tag{3.39}$$

Since $|p(e^{i\theta})| = 1$, $|f(e^{i\theta})| = |g(e^{i\theta})|$; hence, since f belongs to H, so does g. Thus (3.39) is the desired representation of Beurling's theorem.

Finally, to show that p is unique up to a constant factor of modulus 1, suppose that

$$pH = qH \tag{3.40}$$

for functions $p, q \in \mathcal{B}$ which satisfy

$$|p(e^{i\theta})| = 1 = |q(e^{i\theta})|. \tag{3.41}$$

Then by (3.40), there exist $f, g \in H$ such that

$$p = qf, \qquad q = pg,$$

so that by (3.41),

$$1 = |p(e^{i\theta})| = |q(e^{i\theta})f(e^{i\theta})| = |q(e^{i\theta})|\,|f(e^{i\theta})| = |f(e^{i\theta})|$$
$$1 = |q(e^{i\theta})| = |p(e^{i\theta})g(e^{i\theta})| = |p(e^{i\theta})|\,|g(e^{i\theta})| = |g(e^{i\theta})|.$$

By Theorem 3.7 and the maximum principle,

$$|f(0)| \leq 1, \qquad |g(0)| \leq 1. \tag{3.42}$$

Moreover,

$$p = qf = (pg)f = p(gf),$$

so that $1 = gf$. In particular, $1 = g(0)f(0)$. Invoking the maximum principle again, we see from (3.42) that $|f(0)| = 1 = |g(0)|$, so that f and g must both be unimodular constants. □

The elegant proof given above is due to Paul Halmos **[Hal]**.

A function $p \in \mathcal{B}$ such that $|p(e^{i\theta})| = 1$ a.e. on the unit circle is called an *inner function*. As an immediate consequence of Theorem 3.8, we have the following result.

THEOREM 3.9. *Let N be a nontrivial closed subspace of H that is invariant under multiplication by functions in $\mathcal{B}$, i.e., $bN \subset N$ for each $b \in \mathcal{B}$. Then $N = pH$ for some inner function $p \in \mathcal{B}$, which in unique up to a complex constant factor of absolute value* 1.

Of course, each subspace of H of the form pH is invariant under multiplication by $\mathcal{B}$, since $bpH = pbH \subset pH$.

Theorem 3.8 leads to a transparent divisibility theory in the algebra $\mathcal{B}$. We focus on just those aspects of this theory that will be of use to us in the sequel. Our first result concerns divisibility by inner functions.

PROPOSITION 3.10. *An inner function p divides a function b in $\mathcal{B}$ if and only if $pH \supset bH$.*

PROOF. Clearly, if $b = pc$, where c is in B, then $bH = pcH \subset pH$. Conversely, if $bH \subset pH$, then $b = b \cdot 1 = pf$ for some $f \in H$. This shows that $b/p \in H$. But since b is bounded and $|p| = 1$ on the boundary, b/p is bounded on the circle. Therefore, by Theorem 3.7, $b/p \in \mathcal{B}$. □

DEFINITION. Let $a, b \in \mathcal{B}$. Denote by

$$N = \overline{aH + bH} \tag{3.43}$$

the closure of $aH + bH$. According to Theorem 3.8, $N = pH$ for some inner function p, which is designated the greatest common divisor (GCD) of a and b.

This definition is justified by the following

PROPOSITION 3.11. *Let $a, b \in \mathcal{B}$, and let q be an inner function that divides both a and b. Then q divides p, the GCD of a and b.*

PROOF. According to Proposition 3.10, if q divides a, then $aH \subset qH$; similarly if q divides b, then $bH \subset qH$. Therefore $aH + bH \subset qH$. Since qH is closed,

$$\overline{aH + bH} \subset qH.$$

But, by definition, $\overline{aH + bH} = pH$, where p is the GCD of a and b. Since $pH \subset qH$, q divides p. □

DEFINITION. Two functions $a, b \in \mathcal{B}$ are *relatively prime* if their GCD is 1. Thus, according to (3.43), a and b are relatively prime if and only if $aH + bH$ is dense in H.

THEOREM 3.12. *Let a, b and c be functions in $\mathcal{B}$. Suppose that a is relatively prime to both b and c. Then a is relatively prime to their product bc.*

PROOF. By the definition of relatively prime, $aH + bH$ and $aH + cH$ are both dense in H. But then $aH + b(aH + cH) = aH + bcH$ is dense in H. This shows that a and bc are relatively prime. □

Since the upper half plane $\mathbb{H} = \{\operatorname{Im} z > 0\}$ is mapped one-to-one onto the unit disk $\Delta = \{|w| < 1\}$ by the conformal transformation

$$w = \varphi(z) = \frac{z - i}{z + i},$$

Beurling's theorem and the consequent divisibility theory for the algebra $\mathcal{B}$ carry over in a natural fashion to the corresponding spaces of functions defined on $\mathbb{H}$. Specifically, if f is a bounded analytic function on Δ, then $g = f \circ \varphi$ is a bounded analytic function on $\mathbb{H}$, and conversely. This relation is an isometric isomorphism of the algebras $\mathcal{B}(\Delta)$ and $\mathcal{B}(\mathbb{H})$ of bounded analytic functions on Δ and $\mathbb{H}$, respectively, under the sup norms on their respective domains. Functions in $\mathcal{B}(\mathbb{H})$ have boundary values defined a.e. on $\mathbb{R}$. A function $p \in \mathcal{B}(\mathbb{H})$ is an *inner function* if $|p(x)| = 1$ for a.a. $x \in \mathbb{R}$.

THEOREM 3.13. *The only factorizations of e^{iz} as a product of inner functions*

$$e^{iz} = p(z)q(z) \tag{3.44}$$

in the algebra $\mathcal{B}(\mathbb{H})$ of bounded analytic functions in the upper half-plane $\mathbb{H}$ are

$$p(z) = ce^{iaz} \qquad q(z) = c^{-1}e^{ibz},$$

where $a, b \geq 0$, $a + b = 1$, and $|c| = 1$.

PROOF. Assume that (3.44) holds and write $z = x + iy$. Taking the logarithm of the absolute value of (3.44) gives

$$-y = \log|p(z)| + \log|q(z)|. \tag{3.45}$$

Define

$$h(x, y) = -\log|p(z)|. \tag{3.46}$$

Since p and q are assumed to be inner, it follows from (3.45) and (3.46) first that

$$0 \leq h(x, y) \leq y \tag{3.47}$$

and then that h is harmonic on $\mathbb{H}$ and satisfies

$$\lim_{y \to 0} h(x, y) = 0, \qquad x \in \mathbb{R}.$$

Continuing h to the lower half-plane by

$$h(x, y) = -h(x, -y), \tag{3.48}$$

we obtain a function (which we continue to call h) harmonic on both the upper and lower half-planes and continuous on all of $\mathbb{C}$. It follows (cf. [**A**, pp. 172-173]) that h is harmonic on $\mathbb{C}$ and, as is evident from (3.47) and (3.48), that it has at most linear growth. Completing h to an analytic function f on $\mathbb{C}$ and invoking the general version of Liouville's Theorem given in Appendix B, we see that f (and hence h) must be linear. It follows from (3.47) that $h(x, y) = ay$, where $0 \leq a \leq 1$. Thus $p(z) = e^{-ay+iax+id} = e^{id}e^{iaz}$ for some $d \in \mathbb{R}$. It then follows that $q(z) = e^{-by+ibx-id} = e^{-id}e^{ibz}$, where $b = 1 - a$. $\square$

COMMENTS. 1. It can be shown that if $f \in H$ has boundary value function given by (3.30), then

$$\lim_{r \to 1} f(re^{i\theta}) = f(e^{i\theta}) \qquad \text{a.e. on } S^1. \tag{3.49}$$

This is a well-known result of Fatou. One way to prove it is to note that by the Fejér-Lebesgue Theorem [**T**, pp. 415-416], the Fourier series of a function $g \in L^1(S^1)$ is $(C, 1)$ summable, i.e., summable by arithmetic means, to $g(\theta)$ at every point of the Lebesgue set of g (and hence a.e. on S^1). Taking $g(\theta) = f(e^{i\theta})$ as in (3.30) and recalling that $(C, 1)$ summability implies Abel summability [**Har**, p. 108], we obtain (3.49).

2. The divisibility theory in $\mathcal{B}$ discussed above is closely related to the Riesz-Herglotz factorization of functions in the spaces H^p, $1 \leq p \leq \infty$, used by Beurling in his proof of Theorem 3.8. These are the spaces of functions analytic in the unit disk such that

$$\|f\|_p = \lim_{r \to 1} \left(\frac{1}{2\pi} \int_0^{2\pi} |f(re^{i\theta})|^p d\theta \right)^{1/p} < \infty, \quad 1 \leq p < \infty;$$

for $p = \infty$, $H^\infty = \mathcal{B}$, the algebra of bounded analytic functions on the disk with the sup norm. Such functions have L^p boundary functions on the unit circle, all of whose negative Fourier coefficients vanish; and the function on the disk can be recovered as the Cauchy or Poisson integral of its boundary function. (We have had occasion to discuss only the cases $p = 2$ and ∞.)

For such functions, we have the factorization

$$f = cBSF,$$

where

$$B(z) = z^p \prod \left[\frac{\overline{\alpha}_n}{|\alpha_n|} \frac{\alpha_n - z}{1 - \overline{\alpha}_n z} \right]^{p_n}$$

is a (finite or infinite) *Blaschke product* vanishing only at the distinct zeros $\{\alpha_n\}$ (and possibly 0) of f with the corresponding multiplicities p_n (which then satisfy $\sum p_n(1 - |\alpha_n|) < \infty$);

$$S(z) = \exp\left\{ - \int_0^{2\pi} \frac{e^{i\theta} + z}{e^{i\theta} - z} d\mu(\theta) \right\},$$

where μ is a positive measure on the unit circle singular with respect to Lebesgue measure;

$$F(z) = \exp\left\{ \frac{1}{2\pi} \int_0^{2\pi} \log|f(e^{i\theta})| \frac{e^{i\theta} + z}{e^{i\theta} - z} d\theta \right\}$$

is the *outer factor* of f; and c is a unimodular constant. The product $I = BS$ is the *inner factor* of f so that $f = cIF$. See [**D**], [**K**], or [**RR**] for details.

3. For a perspicuous discussion of the relationship between H^p spaces on the disk and on the half-plane, see [**D**, pp. 187-199] or [**RR**, pp. 91-105].

Bibliography

[A] Lars V. Ahlfors, *Complex Analysis*, third edition, McGraw-Hill, 1979.

[B] Arne Beurling, *On two problems concerning linear transformations in Hilbert space*, Acta Math. **81** (1949), 239-255.

[D] Peter L. Duren, *Theory of H^p Spaces*, Academic Press, 1970.

[Hal] Paul Halmos, *Shifts on Hilbert spaces*, J. Reine Angew. Math. **208** (1961), 102-112.

[Har] G.H. Hardy, *Divergent Series*, Oxford University Press, 1949.

[Ho] Kenneth Hoffman, *Banach Spaces of Analytic Functions*, Prentice-Hall, 1962.

[K] Paul Koosis, *Introduction to H_p Spaces*, second edition, Cambridge University Press, 1998.

[L] Peter D. Lax, *Translation invariant spaces*, Acta Math. **101** (1959), 163-178.

[RR] Marvin Rosenblum and James Rovnyak, *Topics in Hardy Classes and Univalent Functions*, Birkhäuser Verlag, 1994.

[T] E.C. Titchmarsh, *The Theory of Functions*, second edition, Oxford University Press, 1939.

3.4. Prediction Theory

1. We denote by X, Y, etc., real-valued square integrable functions defined on some measure space and by (X, Y) the L^2 scalar product.

LEMMA 3.14. *Let $\{X_k\}$ be a countable collection of L^2 functions. Define*

$$e_{jk} = (X_j, X_k). \tag{3.50}$$

Then the matrix (e_{jk}) is symmetric and positive semidefinite.

PROOF. Symmetry is obvious. To say that (e_{jk}) is positive semidefinite means that for any finite set of real numbers u_j,

$$\sum e_{jk}u_ju_k \geq 0.$$

To show this, simply write

$$\sum e_{jk}u_ju_k = \sum (X_j, X_k)u_ju_k = \sum (u_jX_j, u_kX_k) = \|\sum u_\ell X_\ell\|^2 \geq 0.$$

□

A doubly infinite sequence $\{X_k\}$ of L^2 functions is called *stationary* if (3.50) depends only on $j-k$:

$$(X_j, X_k) = e_{j-k}. \tag{3.51}$$

Note that $e_{j-k} = e_{k-j}$ for all j, k.

THEOREM 3.15. *Let $\{X_k\}$ be a stationary sequence of L^2 functions, and suppose that the sequence $\{e_n\}$ defined in* (3.51) *tends to 0 rapidly, say like $O(1/n^2)$. Then*

$$m(\theta) = \sum_{n=-\infty}^{\infty} e_n e^{in\theta} \tag{3.52}$$

is a nonnegative function.

PROOF. Observe that the series in (3.52) converges uniformly and absolutely and hence defines a continuous function on the unit circle S^1. Let g be any smooth function on S^1; write it as the sum of its Fourier series

$$g(\theta) = \sum_{k=-\infty}^{\infty} v_k e^{ik\theta}, \qquad \sum_{k=-\infty}^{\infty} |v_k| < \infty. \tag{3.53}$$

We claim that

$$\int_0^{2\pi} |g(\theta)|^2 m(\theta)d\theta \geq 0. \tag{3.54}$$

Indeed, by the Fourier series representations of g and m, we can write (3.54) as

$$\begin{aligned} \frac{1}{2\pi}\int_0^{2\pi} \sum_{k,\ell,n} v_k e^{ik\theta}\overline{v}_\ell e^{-i\ell\theta} e_n e^{in\theta} d\theta &= \sum_{k-\ell+n=0} e_n v_k \overline{v}_\ell \\ &= \sum_{k,\ell} e_{\ell-k} v_k \overline{v}_\ell = \sum_{k,\ell} e_{k-\ell} v_k \overline{v}_\ell. \end{aligned} \tag{3.55}$$

By Lemma 3.14, the quadratic form $\sum e_{k-\ell}u_ku_\ell$ is positive semidefinite. It follows that the associated Hermitian form, which appears as the right hand side of (3.55) is also. This shows that (3.54) holds as long as only a finite number of the coefficients v_k are nonzero. A routine argument, involving truncation of the sum in (3.55) and the absolute convergence of $\sum v_k$, then yields (3.54) in general.

To complete the proof of Theorem 3.15, observe that any positive smooth function q on S^1 can be written as $|g|^2$, where g is smooth. Thus, it follows from (3.54) that

$$\int_0^{2\pi} q(\theta)m(\theta)d\theta \geq 0$$

for all such q. Clearly, this implies that m is nonnegative. □

2. Suppose now that we are given a stationary sequence X_k and the associated constants e_n. In view of applications (cf. 3.4.3 below), it is natural to ask how well X_0 can be approximated, say in L^2 norm, by linear combinations of the functions X_{-j}, $j = 1, 2, 3, \ldots$. More precisely, how should one choose constants p_j, $j = 1, 2, 3 \ldots$, so that

$$\left\| X_0 - \sum_{j=1}^{\infty} p_j X_{-j} \right\|^2 \tag{3.56}$$

is as small as possible? In formulating this problem, we may as well allow the p_j to take on complex values, since an optimal choice will in any case be real, as is evident from the identity $\|X_0 - (A + iB)\|^2 = \|X_0 - A\|^2 + \|B\|^2$, valid for any real-valued L^2 functions A and B.

Set $p_0 = -1$. Then the quantity to be minimized in (3.56) is

$$\left\| X_0 - \sum_{j=1}^{\infty} p_j X_{-j} \right\|^2 = \left\| \sum_{j=0}^{\infty} p_j X_{-j} \right\|^2 = \sum_{j,k=0}^{\infty} e_{j-k} p_j \overline{p}_k. \tag{3.57}$$

This can be transformed into an extremal problem in complex function theory. Indeed, let

$$p(\theta) = \sum_{k=0}^{\infty} p_k e^{ik\theta}, \tag{3.58}$$

where

$$\sum_{k=0}^{\infty} |p_k|^2 < \infty, \qquad p_0 = -1. \tag{3.59}$$

Then $p \in L^2(S^1)$; and, by the calculation that follows (3.54), we can rewrite (3.57) as

$$\frac{1}{2\pi} \int_0^{2\pi} |p(\theta)|^2 m(\theta) d\theta. \tag{3.60}$$

Now $p(\theta) = f(e^{i\theta})$ almost everywhere, where

$$f(z) = \sum_{k=0}^{\infty} p_k z^k \tag{3.61}$$

is an analytic function on the unit disk of class H^2. The condition $p_0 = -1$ translates into $f(0) = -1$; and the problem of extremizing (3.60) subject to the conditions (3.59) can be restated as that of finding the minimum (or, if the minimum fails to exist, the infimum) of

$$\frac{1}{2\pi} \int_0^{2\pi} |f(e^{i\theta})|^2 m(\theta) d\theta, \tag{3.62}$$

where f ranges over all H^2 functions on the disk which satisfy $f(0) = -1$.

Let us now return to the function m of (3.52). We have already noted that m is a continuous function on S^1 which, by Theorem 3.15, is nonnegative. Assuming for the moment that m is actually positive (and hence bounded away from 0) on S^1, we claim that it can be represented as the square of the absolute value of an H^2 function h on the unit circle:

$$m(\theta) = |h(e^{i\theta})|^2. \tag{3.63}$$

Indeed, evidently $\log m \in L^2$ and so has a Fourier expansion

$$\log m(\theta) = \sum_{k=-\infty}^{\infty} b_k e^{ik\theta}.$$

Since $\log m$ is real valued, $b_{-k} = \overline{b}_k$, so that

$$\log m(\theta) = b(e^{i\theta}) + \overline{b(e^{i\theta})}, \tag{3.64}$$

where

$$b(z) = \frac{1}{2}b_0 + \sum_{k=1}^{\infty} b_k z^k. \tag{3.65}$$

Exponentiating (3.65) and invoking (3.64), we obtain (3.63) with

$$h(z) = e^{b(z)}.$$

Note that

$$h(0) = e^{b(0)} = e^{b_0/2} = \exp\left\{\frac{1}{4\pi}\int_0^{2\pi} \log m(\theta) d\theta\right\}. \tag{3.66}$$

Since m is bounded on S^1, it is clear from (3.63) that h is a bounded analytic function and hence of class H^2. The same holds true for $1/h(z) = \exp\{-b(z)\}$ since m has been assumed to be bounded away from 0.

Using (3.63), we can now rewrite (3.62), the quantity we wish to minimize, as

$$\frac{1}{2\pi}\int_0^{2\pi} |f(e^{i\theta})|^2 |h(e^{i\theta})|^2 d\theta. \tag{3.67}$$

While the function h is determined by the constants e_n, the H^2 function f is arbitrary except for the requirement $f(0) = -1$.

Now

$$\frac{1}{2\pi}\int_0^{2\pi} |f(e^{i\theta})|^2 |h(e^{i\theta})|^2 d\theta = \sum_{k=0}^{\infty} |c_k|^2,$$

where

$$f(z)h(z) = \sum_{k=0}^{\infty} c_k z^k.$$

The coefficient $c_0 = f(0)h(0)$ is fixed; therefore, to minimize (3.67), we make fh constant. Thus

$$f(z)h(z) = f(0)h(0) = -h(0),$$

and we have

$$f(z) = -\frac{h(0)}{h(z)}.$$

The minimum of (3.67) is then $|f(0)h(0)|^2 = |h(0)|^2$. It now follows from (3.66) that the minimum value (3.67), and hence of (3.60), is given by the "geometric mean" of the function m,

$$\exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log m(\theta) d\theta\right\}.$$

It remains to remove the assumption that m is bounded away from 0. Let p be given by (3.58), subject to (3.59). Take $\varepsilon > 0$ and set $m_\varepsilon(\theta) = m(\theta) + \varepsilon$. Then by what has just been shown,

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} |p(\theta)|^2 m(\theta)d\theta &= \frac{1}{2\pi}\int_0^{2\pi} |p(\theta)|^2 m_\varepsilon(\theta)d\theta - \frac{\varepsilon}{2\pi}\int_0^{2\pi} |p(\theta)|^2 d\theta \\ &\geq \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log[m(\theta)+\varepsilon]d\theta\right\} - \frac{\varepsilon}{2\pi}\int_0^{2\pi}|p(\theta)|^2 d\theta.\end{aligned} \tag{3.68}$$

Making $\varepsilon \to 0$ in (3.68), we obtain

$$\frac{1}{2\pi}\int_0^{2\pi} |p(\theta)|^2 m(\theta)d\theta \geq \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log m(\theta)d\theta\right\}$$

by monotone convergence. Thus

$$\inf_p \frac{1}{2\pi}\int_0^{2\pi} |p(\theta)|^2 m(\theta)d\theta \geq \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log m(\theta)d\theta\right\}. \tag{3.69}$$

To prove the opposite inequality, let $\varepsilon > 0$ be fixed. Then there exists a nonvanishing function h_ε in H^2 such that

$$|h_\varepsilon(e^{i\theta})|^2 = m_\varepsilon(\theta), \tag{3.70}$$

$$|h_\varepsilon(0)|^2 = \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log m_\varepsilon(\theta)d\theta\right\} \tag{3.71}$$

and $1/h_\varepsilon$ also belongs to H^2. Set

$$f_\varepsilon(z) = -h_\varepsilon(0)/h_\varepsilon(z); \tag{3.72}$$

then f_ε lies in H^2 and $f_\varepsilon(0) = -1$. Hence by (3.70), (3.71) and (3.72), we have

$$\begin{aligned}\inf_p \frac{1}{2\pi}\int_0^{2\pi} |p(\theta)|^2 m(\theta)d\theta &\leq \frac{1}{2\pi}\int |f_\varepsilon(e^{i\theta})|^2 m(\theta)d\theta \\ &= \frac{1}{2\pi}\int_0^{2\pi} \left|\frac{-h_\varepsilon(0)}{h_\varepsilon(e^{i\theta})}\right|^2 m(\theta)d\theta \\ &= |h_\varepsilon(0)|^2 \frac{1}{2\pi}\int_0^{2\pi}\left|\frac{1}{h_\varepsilon(e^{i\theta})}\right|^2 m(\theta)d\theta \\ &= |h_\varepsilon(0)|^2\frac{1}{2\pi}\int_0^{2\pi}\frac{m(\theta)}{m(\theta)+\varepsilon}d\theta \\ &< |h_\varepsilon(0)|^2 = \exp\left\{\frac{1}{2\pi}\int_0^{2\pi}\log[m(\theta)+\varepsilon]d\theta\right\}.\end{aligned} \tag{3.73}$$

Letting $\varepsilon \to 0$ in (3.73), we obtain

$$\inf_p \frac{1}{2\pi}\int_0^{2\pi} |p(\theta)|^2 m(\theta)d\theta \leq \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log m(\theta)d\theta\right\}. \tag{3.74}$$

Thus, by (3.69) and (3.74),

$$\inf_p \frac{1}{2\pi}\int_0^{2\pi} |p(\theta)|^2 m(\theta)d\theta = \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log m(\theta)d\theta\right\} \tag{3.75}$$

for any nonnegative function m satisfying the conditions stated in Theorem 3.15. In case the integral on the right hand side of (3.75) diverges to $-\infty$, the infimum is 0 (and is not attained).

3. We now give an interpretation of the minimum problem (3.56) in probability theory.

A random medium is sampled at equal time intervals $0, \pm 1, \pm 2, \dots$. Denote by $X(k)$ the random variable that represents sampling at time k. We assume that the $X(k)$ are square integrable. Clearly, they form a stationary sequence. The numbers

$$\mathbb{E}(X(j)X(k)) = e(j-k),$$

where $\mathbb{E}$ is the expected value, are called correlations. The correlations $e(n)$ are known from long time observations of the random medium.

The prediction problem is to predict the present value $X(0)$ from the measured past values of $X(-1), X(-2), \dots$. If we choose a linear predictor

$$\sum_{j=1}^{\infty} p_j X(-j)$$

and call such a predictor optimal if it minimizes the expected value

$$\mathbb{E}((X(0) - \sum_{j=1}^{\infty} p_j X(-j))^2),$$

then we are back at the minimum problem (3.57) posed and solved above.

COMMENTS. 1. The general form of Theorem 3.15, in which the e_n are not necessarily real nor is any rate of decrease assumed, is due to Herglotz [**He**] (cf. [**T**]). It may be stated as

THEOREM 3.15′. *Let $\{e_n\}$ be a doubly infinite sequence such that $e_{-n} = \overline{e}_n$. Then the Hermitian matrix $(e_{jk}) = (e_{j-k})$ is positive semidefinite if and only if*

$$e_n = \int_0^{2\pi} e^{-in\theta} d\mu(\theta)$$

for some nonnegative Borel measure μ on $[0, 2\pi]$.

The continuous version of this result is due to Bochner [**B**].

2. A somewhat more general version of the minimum problem solved above is the following celebrated theorem of Szegő [**Sz**].

THEOREM S. *Let m be a nonnegative integrable function on $[0, 2\pi]$. Then*

$$\inf \frac{1}{2\pi} \int_0^{2\pi} |p(\theta)|^2 m(\theta) d\theta = \exp\left\{ \frac{1}{2\pi} \int_0^{2\pi} \log m(\theta) d\theta \right\},$$

where the infimum is taken over all functions p as in (3.58) *which satisfy* (3.59).

This can be derived from what has already been shown above via an approximation argument; cf. [**DM1**, pp. 191-192].

Twenty years after Szegő obtained this result, Kolmogorov [**K1**], [**K2**] proved a generalization which corresponds to the theorem of Herglotz cited above.

THEOREM K. *Let μ be a (nonnegative) Borel measure on $[0, 2\pi]$ with Radon-Nikodym decomposition $d\mu = \frac{1}{2\pi}m(\theta)d\theta + d\sigma$, where the measure σ is singular with respect to Lebesgue measure. Then*

$$\inf_p \int_0^{2\pi} |p(\theta)|^2 d\mu = \exp\left\{\frac{1}{2\pi}\int_0^{2\pi} \log m(\theta)d\theta\right\},$$

where the infimum is taken over all functions p as in (3.58) which satisfy (3.59).

Proofs of Kolmogorov's Theorem are available in [**Ho**] and [**Koo**]. The applications to prediction theory are due to Kolmogorov, Krein, and Wiener; see the discussion in [**DM2**] and also Wiener's comments [**W**, p. 59].

Bibliography

[B] S. Bochner, *Monotone Funktionen Stieltjessche Integrale und harmonische Analyse*, Math. Ann. **108** (1933), 378-410.

[DM1] H. Dym and H.P. McKean, *Fourier Series and Integrals*, Academic Press, 1971.

[DM2] H. Dym and H.P. McKean, *Gaussian Processes, Function Theory and the Inverse Spectral Problem*, Academic Press, 1976.

[He] G. Herglotz, *Über Potenzreihen mit positivem, reelen Teil im Einheitskreis*, Ber. Verh. Sächs. Akad. Wiss. Leipzig Math.-Natur. Kl. **63** (1911), 501-511; in Gustav Herglotz, Gesammelte Schriften, Vandenhoeck & Ruprecht, 1979, pp. 247-257.

[Ho] Kenneth Hoffman, *Banach Spaces of Analytic Functions*, Prentice-Hall, 1962.

[K1] A.N. Kolmogorov, *Sur l'interpolation et extrapolation des suites stationnaires* C.R. Acad. Sci. Paris **208** (1939), 2043-2045.

[K2] A.N. Kolmogorov, *Stationary sequences in Hilbert space* Bull. Math. Univ. Moscow **2** (no. 6) (1941), 1-40.

[K3] A.N. Kolmogorov, *Interpolation und Extrapolation von stationären zufälligen Folgen*, Bull. Acad. Sci. URSS Sér. Math. **5** (1941), 3-14.

[Koo] Paul Koosis, *Introduction to H_p Spaces*, second edition, Cambridge University Press, 1998.

[Kr1] M. Krein, *On a generalization of some investigations of G. Szegő, V. Smirnoff and A. Kolmogoroff*, C.R. (Doklady) Acad. Sci. URSS (N.S.) **46** (1945), 91-94.

[Kr2] M. Krein, *On a problem of extrapolation of A.N. Kolmogoroff*, C.R. (Doklady) Acad. Sci. URSS (N.S.) **46** (1945), 306-309.

[Sz] G. Szegő, *Beiträge zur Theorie der Toeplitzschen Formen (Erste Mitteilung)*, Math. Z. **6** (1920), 167-202; in Gabor Szegő, Collected Papers, Volume 1, 1915-1927, Birkhäuser, 1982, pp. 237-272, commentary pp. 273-275.

[T] Otto Toeplitz, *Über die Fourier'sche Entwickelung positiver Funktionen*, Rend. Circ. Mat. Palermo **31** (1911), 191-192.

[W] Norbert Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*, MIT Press, 1949.

3.5. The Riesz-Thorin Convexity Theorem

Opinion is unanimous that Marcel Riesz's Convexity Theorem is a deep, important, and powerful tool of modern analysis [**DS**, p. 520], [**Sa**, p. 851] and that G.O. Thorin's proof of this result is a "particularly beautiful instance of the application of complex variable theory to a seemingly unrelated problem in the theory of linear spaces" [**DS**, p. 520]. This section is devoted to the statement, proof, and discussion of this fundamental result.

Let $(M, \mathcal{M}, \mu)$ be a σ-finite measure space, where M is a set, $\mathcal{M}$ the σ-algebra of measurable subsets of M, and μ a (positive) measure defined on $\mathcal{M}$. We denote by $L^p(M)$, $1 \leq p < \infty$, the space of equivalence classes of complex-valued measurable

functions on M satisfying

$$\|f\|_p = \left(\int_M |f(m)|^p d\mu\right)^{1/p} < \infty$$

and by $L^\infty(M)$ the space of equivalence classes of essentially bounded measurable functions on M with the norm

$$\|f\|_\infty = \text{ess sup}_M |f(m)|.$$

Here two functions are considered equivalent if they differ only on a set of μ-measure zero.

In the theorem below, we consider two σ-finite measure spaces, $(U, \mathcal{U}, \mu)$ and $(V, \mathcal{V}, \nu)$, and a linear operator $\mathbf{T}$ mapping the vector space sum $L^{p_0}(U) + L^{p_1}(U)$ into the space of ν-measurable functions on V and mapping $L^{p_j}(U)$ into $L^{q_j}(V)$ for $j = 0, 1$, where $1 \leq p_0, p_1, q_0, q_1 \leq \infty$.

THEOREM 3.16 (Riesz-Thorin). *Suppose that*

$$\mathbf{T} : L^{p_0}(U) \to L^{q_0}(V)$$

and

$$\mathbf{T} : L^{p_1}(U) \to L^{q_1}(V)$$

with norms $M_0 = M(p_0, q_0)$ and $M_1 = M(p_1, q_1)$, respectively. Then for $0 < x < 1$, $\mathbf{T}$ extends to a bounded operator

$$\mathbf{T} : L^p(U) \to L^q(V), \tag{3.76}$$

where

$$\frac{1}{p} = \frac{1-x}{p_0} + \frac{x}{p_1}, \qquad \frac{1}{q} = \frac{1-x}{q_0} + \frac{x}{q_1}, \tag{3.77}$$

with norm $M = M(p, q)$ satisfying

$$M \leq M_0^{1-x} M_1^x. \tag{3.78}$$

REMARK. It follows from the hypothesis of Theorem 3.16 that $\mathbf{T}$ is defined uniquely on $L^{p_0} \cap L^{p_1}$, and it is from this dense subset that it extends to all of L^p. In the original version of this result, proved by Marcel Riesz **[R]** in 1927, it was assumed that $p_0 \leq q_0$ and $p_1 \leq q_1$. The version stated above, as well as the remarkable proof given below, is due to Riesz's student G.O. Thorin **[T1]**, **[T2]**. The full result is now generally known as the Riesz (or Riesz-Thorin) Convexity Theorem, a name based on the fact that it asserts that if $M(p, q)$ is the norm of $\mathbf{T} : L^p(U) \to L^q(V)$, then $\log M(p, q)$ is a convex function of $(1/p, 1/q)$.

PROOF. We fix $0 < x < 1$ and let p and q be as in (3.77). Denote the conjugate index of $1 \leq r \leq \infty$ by r', so that $1/r + 1/r' = 1$, and assume, to begin with, that both p and q' are finite. In order to prove (3.76) and (3.78), we show that the norm of $\mathbf{T}$ when restricted to the simple functions satisfies the inequality (3.78). Since simple functions are dense in $L^p(U)$ for $1 \leq p < \infty$, it then follows that $\mathbf{T}$ extends uniquely to a continuous linear map of $L^p(U)$ into $L^q(V)$; and this extension also satisfies (3.78).

To this end, set

$$\langle h, g \rangle = \int_V h(v) g(v) d\nu, \qquad h \in L^q(V),\ g \in L^{q'}(V)$$

and recall that, by Hölder's inequality,

$$\|h\|_q = \sup_{\|g\|_{q'}=1} |\langle h, g\rangle|. \tag{3.79}$$

Consider the bilinear form $\langle \mathbf{T}f, g\rangle$, where initially f and g are simple functions. We claim that

$$|\langle \mathbf{T}f, g\rangle| \le M_0^{1-x} M_1^x \qquad \text{if} \qquad \|f\|_p = 1, \quad \|g\|_{q'} = 1. \tag{3.80}$$

Suppose this claim has been established. Then, since simple functions are dense in $L^p(U)$ and $L^{q'}(V)$, it follows that (3.80) holds for arbitrary $L^p(U)$, $g \in L^{q'}(V)$ of unit norm. It then follows from (3.79) and (3.80) that

$$M = \sup_{\|f\|_p \le 1} \|\mathbf{T}f\|_q = \sup_{\|f\|_p \le 1} \sup_{\|g\|_{q'} \le 1} |\langle \mathbf{T}f, g\rangle| \le M_0^{1-x} M_1^x,$$

which is (3.78) in case $p < \infty$, $q > 1$.

Turning to the proof of (3.80), let us fix simple functions

$$f = \sum_{j=1}^n a_j \chi_{E_j}, \quad a_j = |a_j| e^{i\theta_j} \ne 0, \quad 1 \le j \le n,$$

$$g = \sum_{k=1}^m b_k \chi_{F_k}, \quad b_k = |b_k| e^{i\varphi_k} \ne 0, \quad 1 \le k \le m$$

such that

$$\begin{aligned} \|f\|_p^p &= \sum_{j=1}^n |a_j|^p \mu(E_j) = 1, \\ \|g\|_{q'}^{q'} &= \sum_{k=1}^n |b_k|^{q'} \nu(F_k) = 1. \end{aligned} \tag{3.81}$$

For $0 \le \operatorname{Re} z \le 1$, put

$$\frac{1}{p(z)} = \frac{1-z}{p_0} + \frac{z}{p_1}, \qquad \frac{1}{q'(z)} = \frac{1-z}{q_0'} + \frac{z}{q_1'}$$

and set

$$\begin{aligned} f_z(u) &= |f(u)|^{p/p(z)} \frac{f(u)}{|f(u)|} = \sum_{j=1}^n |a_j|^{p/p(z)} e^{i\theta_j} \chi_{E_j}(u), \\ g_z(v) &= |g(v)|^{q'/q'(z)} \frac{g(v)}{|g(v)|} = \sum_{k=1}^m |b_k|^{q'/q'(z)} e^{i\varphi_k} \chi_{F_k}(v). \end{aligned} \tag{3.82}$$

Then

$$|f_{iy}(u)| = |f(u)|^{\operatorname{Re}\{p/p(iy)\}} = |f(u)|^{p/p_0},$$
$$|f_{1+iy}(u)| = |f(u)|^{\operatorname{Re}\{p/p(1+iy)\}} = |f(u)|^{p/p_1},$$

so that

$$\begin{aligned} \|f_{iy}\|_{p_0} &= \left(\int_U |f(u)|^p d\mu\right)^{1/p_0} = \|f\|_p^{p/p_0} = 1, \\ \|f_{1+iy}\|_{p_1} &= \left(\int_U |f(u)|^p d\mu\right)^{1/p_1} = \|f\|_p^{p/p_1} = 1. \end{aligned} \tag{3.83}$$

Similarly,

$$\|g_{iy}\|_{q_0'} = 1, \qquad \|g_{1+iy}\|_{q_1'} = 1. \tag{3.84}$$

Now define

$$F(z) = \langle \mathbf{T} f_z, g_z \rangle. \tag{3.85}$$

Then by (3.82) and the linearity of $\mathbf{T}$, we have

$$F(z) = \sum_{k=1}^{m} \sum_{j=1}^{n} A_{jk} |a_j|^{p/p(z)} |b_k|^{q'/q'(z)}, \tag{3.86}$$

where

$$A_{jk} = e^{i(\theta_j + \varphi_k)} \int_V (T\chi_{E_j})(v) \chi_{F_k}(v) d\nu.$$

Each of the summands in (3.86) is an entire function which is bounded in the strip $S = \{z : 0 \le \operatorname{Re} z \le 1\}$; hence F is as well. Now by (3.83) and the definitions of M_0 and M_1,

$$\|\mathbf{T} f_{iy}\|_{q_0} \le M_0 \|f_{iy}\|_{p_0} = M_0$$

and

$$\|\mathbf{T} f_{1+iy}\|_{q_1} \le M_1 \|f_{1+iy}\|_{p_1} = M_1,$$

for all real y. Thus, by Hölder's inequality and (3.84),

$$|F(iy)| = |\langle \mathbf{T} f_{iy}, g_{iy} \rangle| \le \|\mathbf{T} f_{iy}\|_{q_0} \, \|g_{iy}\|_{q_0'} \le M_0$$
$$|F(1+iy)| = |\langle \mathbf{T} f_{1+iy}, g_{1+iy} \rangle| \le \|\mathbf{T} f_{1+iy}\|_{q_1} \, \|g_{1+iy}\|_{q_1'} \le M_1.$$

It now follows from the Three Lines Theorem (Appendix C) that

$$|F(x+iy)| \le M_0^{1-x} M_1^x. \tag{3.87}$$

Since by (3.82)

$$f_x = f \qquad \text{and} \qquad g_x = g,$$

we have from (3.85)

$$F(x) = \langle \mathbf{T} f, g \rangle,$$

which together with (3.87) proves (3.80). This completes the proof of the theorem when $p < \infty$ and $q > 1$.

The remaining cases are easily dealt with. If $p = \infty$ and $q = 1$, then $p_0 = p_1 = \infty$ and $q_0 = q_1 = 1$; so $p(x) = \infty$ and $q(x) = 1$ for all $0 \le x \le 1$, and there is nothing to prove. If $p = \infty$ and $q > 1$, then $p(x) = \infty$ for $0 \le x \le 1$ and $\mathbf{T}$ maps $L^\infty(U)$ into $L^{q_0}(V) \cap L^{q_1}(V)$. In this case, choosing $f_z = f$ for all z allows us to carry out the proof as before. Finally, if $p < \infty$ but $q = 1$, we replace g_z in the proof given above by g and argue as previously. □

REMARK. In many applications, $\mathbf{T}$ is not given initially as a bounded operator on L^{p_0} and L^{p_1} but rather is defined and bounded (in L^{p_0} and L^{p_1} norms) on a dense subset of $L^{p_0} \cap L^{p_1}$. It can then be extended in a unique fashion as a bounded operator on L^{p_0} and L^{p_1}, and the Riesz-Thorin Theorem applies; cf. Section 3.6. In **[FJL]**, an example is given of an operator $\mathbf{T}$ densely defined and bounded in L^{p_0} and L^{p_1} norms which does *not* extend to a bounded operator on L^p for some $p_0 < p < p_1$. In this example, $\mathbf{T}$ is *not* defined on a dense subset of $L^{p_0} \cap L^{p_1}$, and the extensions of $\mathbf{T}$ to L^{p_0} and L^{p_1} actually differ on $L^{p_0} \cap L^{p_1}$. Of course, as follows from the proof above, they also differ on the class of simple functions. We thank Michael Cwikel for having brought this example to our attention.

As an initial illustration of the power of Riesz's Convexity Theorem, we have the following painless proof of Young's Inequality in the theory of convolution operators on the spaces $L^p(\mathbb{R})$. Recall that the convolution $f * g$ of f and g is defined by

$$(f * g)(x) = \int_{-\infty}^{\infty} f(x-t)g(t)dt.$$

Young's Inequality asserts that for $1 \le p, r \le \infty$,

$$\|f * g\|_s \le \|f\|_r \|g\|_p, \tag{3.88}$$

where

$$\frac{1}{p} + \frac{1}{r} = 1 + \frac{1}{s}. \tag{3.89}$$

For the proof, observe first that for fixed $f \in L^1$, a simple calculation involving Fubini's Theorem shows that the operator $\mathbf{T}$ defined by

$$\mathbf{T}g = f * g$$

maps L^1 boundedly into L^1 with norm bounded by (actually, equal to) $\|f\|_1$. Since $\mathbf{T}$ trivially maps L^∞ into L^∞ with the same bound, Riesz's theorem with $(p_0, q_0) = (1,1)$ and $(p_1, q_1) = (\infty, \infty)$ yields

$$\|f * g\|_p \le \|f\|_1 \|g\|_p, \tag{3.90}$$

which is (3.88) with $s = p$ and $r = 1$. Now fix $g \in L^p$ and define $\mathbf{S}$ by

$$\mathbf{S}f = f * g.$$

Then (3.90) shows that

$$\mathbf{S} : L^1 \to L^p \qquad \text{with} \qquad \|\mathbf{S}\| \le \|g\|_p, \tag{3.91}$$

while Hölder's inequality gives

$$\mathbf{S} : L^q \to L^\infty \qquad \text{with} \qquad \|\mathbf{S}\| \le \|g\|_q \tag{3.92}$$

when

$$\frac{1}{p} + \frac{1}{q} = 1. \tag{3.93}$$

Plugging (3.91), (3.92) and (3.93) into Riesz's theorem, we obtain (3.88) where r and s are related by (3.89).

Another application, more striking yet, is the following proof of the Hausdorff-Young Theorem. Let $\mathbf{T}$ be the operator that maps an integrable function on the unit circle to its sequence of Fourier coefficients

$$\hat{f}(n) = \frac{1}{2\pi} \int_0^{2\pi} f(\theta) e^{-in\theta} d\theta, \quad n \in \mathbb{Z}.$$

By Parseval's Theorem,

$$\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(\theta)|^2 d\theta = \|f\|_2^2,$$

so that $\mathbf{T} : L^2 \to \ell^2$ with norm 1. On the other hand, it is evident that $\mathbf{T} : L^1 \to \ell^\infty$ with norm 1. Taking $(p_0, q_0) = (1, \infty)$ and $(p_1, q_1) = (2,2)$ in Riesz's theorem, we see that for $1 < p < 2$ and $1/p + 1/q = 1$,

$$\mathbf{T} : L^p \to \ell^q \qquad \text{with} \qquad \|\mathbf{T}\| = 1,$$

or

$$\left(\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^q\right)^{1/q} \leq \left(\frac{1}{2\pi}\int_0^{2\pi} |f(\theta)|^p d\theta\right)^{1/p}.$$

This is the Hausdorff-Young Theorem.

Of this last argument, Littlewood writes, "**T** thus produces a high-brow result 'out of nothing'; we experience something like the intoxication of the early days of projecting conics into circles" [**L**, p. 41].

COMMENTS. 1. An extensive discussion of theorems concerning bilinear and multilinear forms, including Riesz's original proof of his convexity theorem, is in [**HLP**, Chapter 8].

2. Mention should also be made of an extension of the Riesz-Thorin Theorem due to E.M. Stein [**St**, Theorem 1]; cf. [**SW**, pp. 205-209]. Here, instead of a single operator T, one has an analytic family of operators T_z, $0 \leq \operatorname{Re} z \leq 1$, such that T_z is bounded from L^{p_j} to L^{q_j} for $\operatorname{Re} z = j$, $j = 0, 1$, and satisfies an appropriate growth condition as $\operatorname{Im} z \to \pm\infty$. The conclusion then is that for $0 < x < 1$, $T_x : L^p \to L^q$ boundedly, where p and q are defined by (3.77).

3. Thorin's essential insight was elaborated into the Complex Method of Interpolation by Calderón [**C**]; cf. [**BL**, Chapter 4]. This approach to the Riesz-Thorin Theorem is followed in [**K**, pp. 117-121].

4. The Hausdorff-Young Theorem has a companion result, which also follows instantly from the Riesz Convexity Theorem. Specifically, let $1 \leq p \leq 2$ and suppose $1/p + 1/q = 1$. Then if $\{a_n\}_{n=-\infty}^{\infty} \in \ell^p$, there exists $f \in L^q(S^1)$ such that $\hat{f}(n) = a_n$; moreover, $\|f\|_q \leq \left(\sum_{n=-\infty}^{\infty} |a_n|^p\right)^{1/p}$. For the proof, simply note that if $\{a_n\} \in \ell^1$, then

$$f(\theta) = \sum_{n=-\infty}^{\infty} a_n e^{in\theta} \tag{3.94}$$

is continuous on S^1 and $\hat{f}(n) = a_n$. Clearly $\|f\|_\infty \leq \|\{a_n\}\|_1$; i.e., the map $\mathbf{T} : \ell^1 \to L^\infty$ defined by (3.94) has norm 1. As before, (3.94) also defines a map $\mathbf{T} : \ell^2 \to L^2$ having norm 1. Thus, interpolating between $(1, \infty)$ and $(2, 2)$, we see that (3.94) defines a continuous map $T : \ell^p \to L^q$ such that $\|f\|_q \leq \left(\sum_{n=-\infty}^{\infty} |a_n|^p\right)^{1/p}$.

5. For more on Olof Thorin, whose entire professional career was spent working for an insurance company, see [**BoGP**].

Bibliography

[BL] Jöran Bergh and Jörgen Löfström, *Interpolation Spaces*, Springer-Verlag, 1976.

[BoGP] Lennart Bondesson, Jan Grandell, and Jaak Peetre, *The life and work of Olof Thorin (1912-2004)*, Proc. Est. Acad. Sci. **57** (2008), 18-25.

[C] A.P. Calderón, *Intermediate spaces and interpolation, the complex method*, Studia Math. **24** (1964), 113-190.

[DS] Nelson Dunford and Jacob T. Schwartz, *Linear Operators Part 1: General Theory*, Wiley Interscience, 1957.

[FJL] E.B. Fabes, Max Jodeit, Jr., and J.E. Lewis, *On the spectra of a Hardy kernel*, J. Funct. Anal. **21** (1976), 187-194.

[HLP] G.H. Hardy, J.E. Littlewood, and G. Pólya, *Inequalities*, Cambridge University Press, 1959.

[K] Yitzhak Katznelson, *An Introduction to Harmonic Analysis*, third edition, Cambridge University Press, 2004.

[L] J.E. Littlewood, *Littlewood's Miscellany*, edited by Béla Bollobás, Cambridge University Press, 1986.

[R] Marcel Riesz, *Sur les maxima des formes bilinéaires et sur les fonctionnelles linéaires*, Acta Math. **49** (1927), 465-497.

[Sa] R. Salem, *Convexity theorems*, Bull. Amer. Math. Soc. **55** (1949), 851-860.

[St] Elias M. Stein, *Interpolation of linear operators*, Trans. Amer. Math. Soc. **83** (1956), 482-492.

[SW] Elias M. Stein and Guido Weiss, *Introduction to Fourier Analysis on Euclidean Spaces*, Princeton University Press, 1971.

[T1] G.O. Thorin, *An extension of a convexity theorem due to M. Riesz*, Kungl. Fysiogr. Sällsk. i Lund Förh. **8** (1938), 166-170.

[T2] G.O. Thorin, *Convexity theorems generalizing those of M. Riesz and Hadamard with some applications*, Medd. Lunds Univ. Mat. Sem. **9** (1948), 1-58.

3.6. The Hilbert Transform

Let h be a real-valued integrable function on $\mathbb{R}$. The Cauchy integral

$$f(z) = \frac{1}{\pi i}\int_{\mathbb{R}} \frac{h(t)}{t-z}\,dt \tag{3.95}$$

defines a function $f(z)$, or rather two functions, one analytic on the upper half-plane, the other in the lower half-plane. We restrict z to the upper half plane.

Writing $z = x + iy$, we can express the real and imaginary parts of f as follows:

$$\begin{aligned} f(z) &= \frac{1}{\pi i}\int_{\mathbb{R}} \frac{h(t)(t-\overline{z})}{|t-z|^2}\,dt \\ &= \frac{1}{\pi}\int_{\mathbb{R}} h(t)\frac{y}{(x-t)^2+y^2}\,dt + \frac{i}{\pi}\int_{\mathbb{R}} h(t)\frac{x-t}{(x-t)^2+y^2}\,dt. \end{aligned} \tag{3.96}$$

THEOREM 3.17. *Suppose that h is a real-valued continuously differentiable function of compact support, and let f be given by (3.95). Then*

(i) *As $|z| \to \infty$,*

$$f(z) = O(1/|z|); \tag{3.97}$$

(ii) *f extends continuously from the upper half-plane to the real axis, and*

$$f(x) = h(x) + ik(x), \tag{3.98}$$

where

$$k(x) = \lim_{\varepsilon\to 0}\frac{1}{\pi}\int_{|x-t|>\varepsilon} \frac{h(t)}{x-t}\,dt. \tag{3.99}$$

REMARKS. 1. The right hand side of (3.99) is often written (and we shall write it) as

$$\mathrm{PV}\frac{1}{\pi}\int \frac{h(t)}{x-t}\,dt.$$

Here PV stands for "principal value".

2. Theorem 3.17 holds under significantly weaker hypotheses than we have stated. However, the version given above is adequate for our purposes, as $C_c^1(\mathbb{R})$ is dense in $L^p(\mathbb{R})$ for $1 \le p < \infty$.

PROOF. (i) As $z \to \infty$, $z/(t-z) \to -1$ uniformly on the (compact) support of h; hence

$$\lim_{z\to\infty} |zf(z)| = \lim_{z\to\infty} \left| \frac{1}{\pi i} \int_{\mathbb{R}} h(t) \frac{z}{t-z} dt \right| = \left| \frac{1}{\pi} \int_{\mathbb{R}} h(t) dt \right| < \infty.$$

(ii) By (3.96), there are two claims to prove. The first of these is that

$$\lim_{y\to 0} \frac{1}{\pi} \int_{\mathbb{R}} h(t) \frac{y}{(x-t)^2+y^2} dt = h(x), \qquad x \in \mathbb{R}. \tag{3.100}$$

This follows in routine fashion from the fact that

$$P_y(x) = \frac{1}{\pi} \frac{y}{x^2+y^2}$$

is an approximate identity in the sense that

(a) $P_y(x) > 0$ for all $x \in \mathbb{R}$, $y > 0$;

(b) $\int_{\mathbb{R}} P_y(x) dx = 1$ for all $y > 0$; and

(c) $\lim_{y\to 0} \int_{-\varepsilon}^{\varepsilon} P_y(x) dx = 1$ for each $\varepsilon > 0$.

The integral on the left hand side of (3.100) is the convolution of h with the kernel P_y, and one sees easily that this converges uniformly (in x) to h as $y \to 0$.

The second assertion of (ii) is that

$$\lim_{y\to 0} \frac{1}{\pi} \int_{\mathbb{R}} h(t) \frac{x-t}{(x-t)^2+y^2} dt = \mathrm{PV}\, \frac{1}{\pi} \int \frac{h(t)}{x-t} dt, \tag{3.101}$$

where the right hand side exists and is a continuous function. To this end, let us first note that

$$\begin{aligned} \mathrm{PV}\, \frac{1}{\pi} \int_{\mathbb{R}} \frac{h(t)}{x-t} dt &= \lim_{\varepsilon\to\infty} \left[\frac{1}{\pi} \int_{-\infty}^{x-\varepsilon} \frac{h(t)}{x-t} dt + \frac{1}{\pi} \int_{x+\varepsilon}^{\infty} \frac{h(t)}{x-t} dt \right] \\ &= \lim_{\varepsilon\to 0} \frac{1}{\pi} \int_{\varepsilon}^{\infty} \frac{h(x-t)-h(x+t)}{t} dt \\ &= \frac{1}{\pi} \int_0^{\infty} \frac{h(x-t)-h(x+t)}{t} dt \end{aligned} \tag{3.102}$$

since

$$\left| \frac{h(x-t)-h(x+t)}{t} \right| \le 2 \max_{\mathbb{R}} |h'(s)| < \infty. \tag{3.103}$$

Set

$$k(x+iy) = \frac{1}{\pi} \int_{-\infty}^{\infty} h(t) \frac{x-t}{(x-t)^2+y^2} dt.$$

We rewrite this as

$$k(x+iy) = \frac{1}{\pi} \int_0^{\infty} [h(x-t)-h(x+t)] \frac{t}{t^2+y^2} dt. \tag{3.104}$$

Hence, denoting the common value of the left and right hand sides of (3.102) by $k(x)$, we have from (3.104) and (3.103)

$$\begin{aligned}|k(x+iy)-k(x)| &\le \frac{1}{\pi}\int_0^\infty |h(x-t)-h(x+t)|\left|\frac{t}{t^2+y^2}-\frac{1}{t}\right| dt \\ &= \frac{1}{\pi}\int_0^\infty \frac{|h(x-t)-h(x+t)|}{t}\,\frac{y^2}{t^2+y^2}dt \\ &\le 2\max_{\mathbb{R}}|h'(s)|\frac{1}{\pi}\int_0^\infty \frac{y^2}{t^2+y^2}dt \\ &= y\max_{\mathbb{R}}|h'(s)|\end{aligned}$$

for all $x \in \mathbb{R}$. Thus

$$\lim_{y\to 0} k(x+iy) = k(x) \qquad \text{uniformly on } \mathbb{R},$$

which proves (3.101). Since $k(x+iy)$ is continuous as a function of x for each $y > 0$, it follows that $k(x)$ is also continuous. This completes the proof. □

For $h \in C_c^1(\mathbb{R})$, we define the *Hilbert transform* of h by

$$\mathbf{H}h(x) = \mathrm{PV}\,\frac{1}{\pi}\int \frac{h(t)}{x-t}dt;$$

by what we have shown, it relates the real to the imaginary part of the boundary values of analytic functions in the upper half plane satisfying (3.97).

THEOREM 3.18. *The Hilbert transform extends to an isometry of $L^2(\mathbb{R}) \to L^2(\mathbb{R})$.*

PROOF. Take $h \in C_c^1(\mathbb{R})$ and let f be defined by (3.95). Now f^2 is analytic in $\operatorname{Im} z > 0$; hence, by Cauchy's Theorem,

$$\int_\Gamma [f(z)]^2 dz = 0 \tag{3.105}$$

for any closed contour Γ in the upper half-plane. Take Γ to consist of the line segment $x + i\varepsilon$, $-R \le x \le R$ and the semicircle $z = Re^{i\theta} + i\varepsilon$, $0 \le \theta \le \pi$. Now let $\varepsilon \to 0$ and $R \to \infty$. It follows from (3.97) that the integral over the semicircle tends to 0 as $R \to \infty$, while it follows from (3.97) and (3.98) that the integral over the segment tends to

$$\int_{\mathbb{R}} (h+ik)^2 dx = 0. \tag{3.106}$$

Taking the real part of (3.106) gives

$$\int_{\mathbb{R}} h^2 dx = \int_{\mathbb{R}} k^2 dx.$$

Thus, $\mathbf{H}$ is an isometry in L^2 norm when restricted to $C_c^1(\mathbb{R})$. Since $C_c^1(\mathbb{R})$ is dense in $L^2(\mathbb{R})$, there is a unique extension of $\mathbf{H}$ as a continuous linear map $\mathbf{H} : L^2(\mathbb{R}) \to L^2(\mathbb{R})$; and this map is isometric, i.e., $\|\mathbf{H}h\|_2 = \|h\|_2$ for all $h \in L^2(\mathbb{R})$. □

More generally, we have

THEOREM 3.19. *The Hilbert transform extends to a bounded map $\mathbf{H} : L^p(\mathbb{R}) \to L^p(\mathbb{R})$ for all p, $1 < p < \infty$.*

PROOF. Suppose $p = 2m$, m an integer, and consider the analytic function f^{2m}. By Cauchy's Theorem,

$$\int_\Gamma [f(z)]^{2m} dz = 0$$

for any closed contour in the upper half-plane. We choose the same contour as in (3.105) and let $\varepsilon \to 0$, $R \to \infty$, to obtain

$$\int_{\mathbb{R}} (h + ik)^{2m} dx = 0.$$

The real part of this relation is

$$\int_{\mathbb{R}} \sum_{j=0}^{m} (-1)^j \binom{2m}{2j} h^{2m-2j} k^{2j} dx = 0,$$

so clearly,

$$\int_{\mathbb{R}} k^{2m} dx \le \int_{\mathbb{R}} \sum_{j=0}^{m-1} \binom{2m}{2j} h^{2m-2j} k^{2j} dx. \tag{3.107}$$

For each $\varepsilon > 0$, there exists $C(\varepsilon) > 0$ such that

$$h^{2m-2j} k^{2j} \le C(\varepsilon) h^{2m} + \varepsilon k^{2m} \qquad \text{for all} \quad 1 < j < m. \tag{3.108}$$

Now

$$\sum_{j=0}^{m} \binom{2m}{2j} = 2^{2m-1},$$

so combining (3.107) and (3.108) and choosing $\varepsilon < 1/2^{2m-1}$ gives

$$\int_{\mathbb{R}} k^{2m} dx \le A \int_{\mathbb{R}} h^{2m} dx$$

for an appropriately large value of $A > 0$. Thus $\mathbf{H}$ is a bounded map from L^{2m} to L^{2m}. It then follows from the Riesz Convexity Theorem that $\mathbf{H}$ is a bounded map from L^p to L^p, $2 \le p \le 2m$. Since m is arbitrary, $\mathbf{H} : L^p \to L^p$ is bounded for all $2 \le p < \infty$.

To complete the proof for $1 < p < 2$, we use some standard facts from functional analysis. For a Banach space X and a continuous linear functional x^* in the dual space X^*, we write (x, x^*) for $x^*(x)$, where $x \in X$. Recall now that if $\mathbf{T}$ is a bounded linear map between Banach spaces $\mathbf{T} : X \to Y$, its adjoint (or transpose) $\mathbf{T}^* : Y^* \to X^*$, defined by

$$(\mathbf{T}x, y^*) = (x, \mathbf{T}^* y^*), \qquad y^* \in Y^*,$$

is also a bounded linear map and the operator norms $\|\mathbf{T}\|$ and $\|\mathbf{T}^*\|$ coincide [**FA**, p. 163]. Now

$$\mathbf{H} : L^p \to L^p, \qquad 2 \le p < \infty,$$

so $\mathbf{H}^* : (L^p)^* \to (L^p)^*$ and $\|\mathbf{H}\| = \|\mathbf{H}^*\|$. But it is well-known [**FA**, p. 79] that $(L^p)^* = L^{p'}$, where $1/p + 1/p' = 1$. Moreover, a simple calculation based on the definition of the adjoint shows that $\mathbf{H}^* = -\mathbf{H}$. It follows that the norm of $\mathbf{H} : L^{p'} \to L^{p'}$ equals the norm of $\mathbf{H} : L^p \to L^p$. Since the latter were shown to be finite for $2 < p < \infty$, it follows that they are bounded for $1 < p' < 2$ as well. This completes the proof. □

COMMENT. The norm of the Hilbert transform on $L^p(\mathbb{R})$, i.e., the smallest constant A_p such that $\|\mathbf{H}h\|_p \leq A_p\|h\|_p$ for all $h \in L^p(\mathbb{R})$, is given by

$$A_p = \begin{cases} \tan \pi/2p & 1 < p \leq 2 \\ \cot \pi/2p & 2 \leq p < \infty. \end{cases}$$

This was conjectured by Gohberg and Krupnik [**GK**], who proved it for $p = 2^n$ $(n = 1, 2, \dots)$, and proved in full generality by Pichorides [**P**].

Bibliography

[GK] I.C. Gohberg and N.Ja. Krupnik, *On the norm of the Hilbert transform in the space L^p*, Funct. Anal. Appl. **2** (1968), 180-181.

[P] S.K. Pichorides, *On the best values of the constants in the theorems of M. Riesz, Zygmund and Kolmogorov*, Studia Math. **44** (1972), 165-179.

CHAPTER 4

Harmonic Analysis

It has been said that the three most effective problem-solving devices in mathematics are calculus, complex variables, and the Fourier transform. In this chapter, we explore some of the relations between these latter two in order to illustrate what Arne Beurling has called "the close relation between analytic functions and harmonic analysis on Euclidean groups."

4.1. Fourier Uniqueness via Complex Variables (d'après D.J. Newman)

The uniqueness theorem for the one-dimensional Fourier transform asserts that if $f \in L^1(\mathbb{R})$ and

$$\mathcal{F}f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(t) e^{ixt} dt = \hat{f}(x)$$

vanishes identically for $x \in \mathbb{R}$, then $f = 0$, i.e., $f(t) = 0$ a.e. The following proof of this result, due to Donald Newman [**N**], is a true *tour de force* of complex variables.

PROOF. Suppose $\hat{f}(x) = 0$ for all $x \in \mathbb{R}$. Fix $a \in \mathbb{R}$ and denote by $F_a(x)$ the common value of both sides of

$$\int_{-\infty}^{a} f(t) e^{ix(t-a)} dt = -\int_{a}^{\infty} f(t) e^{ix(t-a)} dt. \tag{4.1}$$

Now let x take on complex values; the integral on the left side of (4.1) then defines a bounded, continuous function on $\{\operatorname{Im} x \leq 0\}$ which is analytic on $\mathbb{H}_- = \{\operatorname{Im} x < 0\}$, while the integral on the right is continuous and bounded on $\{\operatorname{Im} x \geq 0\}$ and analytic on $\mathbb{H} = \{\operatorname{Im} x > 0\}$. Thus F_a is continuous on $\mathbb{H}_- \cup \mathbb{R} \cup \mathbb{H} = \mathbb{C}$ and analytic on $\mathbb{H}_- \cup \mathbb{H}$, and hence by Morera's Theorem analytic throughout $\mathbb{C}$. But F_a is also bounded on $\mathbb{C}$, so by Liouville's Theorem it is constant. Taking $x = is$ $(s > 0)$ in the right hand side of (4.1) and letting $s \to +\infty$ shows that this constant is 0. Thus

$$0 = F_a(0) = \int_{-\infty}^{a} f(t) dt, \tag{4.2}$$

and this holds for each $a \in \mathbb{R}$. Differentiating (4.2) yields $f(a) = 0$ a.e. $\square$

Bibliography

[N] D.J. Newman, *Fourier uniqueness via complex variables*, Amer. Math. Monthly **81** (1974), 379-380.

4.2. A Curious Functional Equation

Every linear function $f(x) = mx$ satisfies the functional equation

$$f(x+y) = f(x) + f(y). \tag{4.3}$$

Conversely, every function f that satisfies (4.3) and is continuous at $x = 0$ is linear; this is a classical result.

Set $y = x$ in (4.3):

$$f(2x) = 2f(x). \tag{4.4}$$

THEOREM 4.1. *Every solution of (4.4) that is once differentiable at $x = 0$ is linear.*

PROOF. Setting $x = 0$ in (4.4) shows that $f(0) = 0$. Applying (4.4) n times gives

$$f(x) = 2^n f(x/2^n) \tag{4.5}$$

Since f is differentiable at $x = 0$,

$$f(y) = my + \varepsilon y, \tag{4.6}$$

where $m = f'(0)$, and $\varepsilon = \varepsilon(y)$ tends to zero as y tends to zero.

Set $y = x/2^n$ into (4.6), and use (4.5):

$$f(x) = 2^n(mx/2^n + \varepsilon x/2^n) = mx + \varepsilon x.$$

As n tends to ∞, ε tends to zero, giving $f(x) = mx$. □

The condition that f be differentiable at $x = 0$ cannot be replaced by requiring mere Lipschitz continuity. Indeed, every function of the form

$$f(x) = x^p, \tag{4.7}$$

where

$$p = 1 + 2\pi i n/\log 2, \qquad n \text{ an integer}, \tag{4.8}$$

satisfies equation (4.4); these functions are all Lipschitz continuous at $x = 0$.

We now turn to a continuous analogue of equation (4.4):

$$\frac{1}{x}\int_0^x f(y)dy = f(x/2). \tag{4.9}$$

Clearly, all functions of the form

$$f(x) = c + mx$$

satisfy (4.8); what additional restrictions characterize these solutions?

Each function of form (4.7) satisfies equation (4.9), provided that the exponent p satisfies the relation

$$2^p = p + 1. \tag{4.10}$$

This transcendental equation has infinitely many solutions $p_1, p_2, \ldots$, where $|p_n|$ tends to infinity as n does. Equation (4.10) implies that $|2^{p_n}|$ tends to infinity, from which it follows that $\operatorname{Re} p_n$ tends to infinity. When $\operatorname{Re} p_n > N$, $f_n(x) = x^{p_n}$ is N times differentiable at the origin. This shows that requiring f to have N derivatives at $x = 0$ does not single out $f(x) = c + mx$ as the only solution of (4.9).

To investigate this question further, multiply equation (4.9) by x and differentiate; we get

$$f(x) = f(x/2) + \frac{1}{2}xf'(x/2). \tag{4.11}$$

Differentiating (4.11) n times with respect to x gives

$$f^{(n)}(x) = a_n f^{(n)}(x/2) + b_n x f^{(n+1)}(x/2), \tag{4.12}$$

where the sequences a_n, b_n satisfy the recursions

$$a_{n+1} = \frac{1}{2}a_n + b_n, \qquad b_{n+1} = \frac{1}{2}b_n.$$

It follows that $b_n = (1/2)^n$ and

$$a_0 = 1,\ a_1 = 1,\ a_2 = 3/4, \dots .$$

Thus $a_n < 1$ for $n > 1$.

Assume that f is C^∞ at $x = 0$, and set $x = 0$ in equation (4.12). We get

$$f^{(n)}(0) = a_n f^{(n)}(0).$$

Since $a_n < 1$ for $n > 1$, we can conclude that

$$f^{(n)}(0) = 0, \qquad n > 1. \tag{4.13}$$

From this we deduce

THEOREM 4.2. *If a solution f of (4.9) is analytic at $x = 0$, then $f(x) = c+mx$.*

Our aim is the following less obvious result.

THEOREM 4.3. *A solution f of (4.9) which is infinitely differentiable at $x = 0$ is of the form $f(x) = c + mx$.*

PROOF. According to (4.13), all derivatives of order > 1 of such a function f are zero at $x = 0$. Subtracting $c + mx$ from f, where $c = f(0)$, $m = f'(0)$, gives a function, which we continue to denote by f, which vanishes along with all its derivatives at the origin.

We change variables $x = e^s$ and write

$$f(e^s) = g(s). \tag{4.14}$$

Set $a = \log 2$; then

$$x/2 = e^{s-a}, \qquad e^a = 2.$$

Denote d/ds by a dot; differentiating

$$f(x/2) = g(s-a)$$

with respect to s gives

$$\frac{1}{2}xf'(x/2) = \dot{g}(s-a).$$

Setting this into (4.11), we get

$$g(s) = g(s-a) + \dot{g}(s-a). \tag{4.15}$$

We take the Fourier transform of (4.15) over $(-\infty, a)$

$$\int_{-\infty}^{a} g(s)e^{isz}ds = \int_{-\infty}^{a} g(s-a)e^{isz}ds + \int_{-\infty}^{a} \dot{g}(s-a)e^{isz}ds. \tag{4.16}$$

Introducing $s - a = r$ as new variable of integration on the right in (4.16) and integrating the second term by parts gives

$$e^{iaz}\int_{-\infty}^{0} g(r)e^{irz}dr - ize^{iaz}\int_{-\infty}^{0} g(r)e^{irz}dr + e^{iaz}g(0).$$

Set

$$\int_{-\infty}^{0} g(r)e^{irz}dr = G(z) \tag{4.17}$$

and

$$\int_{0}^{a} g(s)e^{isz}ds - e^{iaz}g(0) = R(z). \tag{4.18}$$

Then (4.16) can be rewritten as

$$R(z) = D(z)G(z), \tag{4.19}$$

where

$$D(z) = e^{iaz}(1 - iz) - 1. \tag{4.20}$$

From the definition (4.18) of $R(z)$, we see that $R(z)$ is an entire function which is bounded in the upper half plane:

$$|R(z)| \leq \text{const.} \qquad \text{for} \quad \operatorname{Im} z \geq 0. \tag{4.21}$$

Since f has a zero of infinite order at $x = 0$,

$$f(x) = O(x^n), \qquad n > 0.$$

From equation (4.14) relating f and g, we conclude that for any n,

$$g(s) = O(e^{ns}) \qquad \text{as} \quad s \to -\infty.$$

Thus the integral (4.17) converges for all complex values of z, and it follows that $G(z)$ is an *entire* function of z.

We rewrite equation (4.19) as

$$\frac{R(z)}{D(z)} = G(z). \tag{4.22}$$

Since G is entire, the zeros of $D(z)$ are matched by the zeros of $R(z)$. The zeros of $D(z)$ are of the form $z = ip_n$, where p_n is a root of (4.10).

LEMMA 4.4. *$G(z)$ is bounded in the upper half plane* $\operatorname{Im} z > 0$.

PROOF. We first estimate $D(z)$ from below; we claim that

$$|D(z)| \geq 1 \tag{4.23}$$

on the rays

$$z = \pi n/a + iv, \tag{4.24}$$

n any odd integer, $v \geq 0$. To see this, set (4.24) into the definition (4.20) of $D(z)$:

$$D(\pi n/a + iv) = -e^{-av}(1 + v - i\pi n/a) - 1.$$

It follows that $\operatorname{Re} D(\pi n/a + iv) < -1$, from which (4.23) follows.

Next we show

$$|D(z)| > 1/2 \tag{4.25}$$

on the boundary of the rectangle

$$\pi n/a \le \operatorname{Re} z \le \pi n/a + 2, \qquad 0 \le \operatorname{Im} z \le k, \tag{4.26}$$

where n is an odd integer $\neq -3, -1, 1$, and k is sufficiently large.

The estimate (4.25) follows from (4.23) on the vertical side of the rectangles. On the top, where $\operatorname{Im} z = k$, the exponential factor e^{iaz} in (4.20) is exponentially small, so (4.25) follows. At the bottom $\operatorname{Im} z = 0$, the exponential factor e^{iaz} in (4.20) has absolute value 1 and $|iz| \ge 3$, so again (4.25) follows.

We have pointed out that $|R(z)|$ is bounded for $\operatorname{Im} z > 0$. Therefore, by (4.25), $|R(z)/D(z)|$ is bounded by twice that constant on the boundary of the rectangles (4.26). But since $|R(z)/D(z)|$ is analytic, it follows from the maximum principle that $|R(z)/D(z)|$ is bounded by the same constant inside the rectangle. Letting k tend to infinity, we conclude that

$$|R(z)/D(z)| \le \text{const.}$$

for all z in the upper half plane except for the strip $|\operatorname{Re} z| \le 3$. Using the same argument as before, we can show that $R(z)/D(z)$ is bounded in the strip $|\operatorname{Re} z| \le 3$, $\operatorname{Im} z \ge k$ as well. Since the remaining portion of the upper half plane is compact, and $R(z)/D(z)$ is analytic, it is bounded there as well.

This completes the proof of Lemma 4.4. □

Now $G(z)$ is defined in (4.17) as the Fourier transform of $g(x)$ on $(-\infty, 0]$ and hence is bounded in the lower half plane $\operatorname{Im} z < 0$. According to Lemma 4.4, $G(z)$ is uniformly bounded in the upper half plane $\operatorname{Im} z > 0$ as well. Thus the entire function $G(z)$ is bounded in the whole complex plane and hence constant by Liouville's Theorem. Since $G(-iy) \to 0$ as $y \to +\infty$, G vanishes identically; so by Fourier uniqueness, it follows that $g(s) = 0$ for $s \le 0$. Equation (4.14) shows that

$$f(x) = g(\log x);$$

therefore, $f(x) = 0$ for $0 < x < 1$. Now we use the functional equation (4.11) to conclude inductively that $f(x) = 0$ for $x < 2^n$, $n = 1, 2, \ldots$. This shows that $f(x) = 0$ for all $x > 0$, as asserted in Theorem 4.3. □

Bibliography

[L] Peter D. Lax, *A curious functional equation*, J. Anal. Math. **105** (2008), 383-389.

4.3. Uniqueness and Nonuniqueness for the Radon Transform

1. Let f be defined on $\mathbb{R}^2$ and suppose that

$$\int_\ell f\, ds = 0 \tag{4.27}$$

for each line ℓ. Must f vanish almost everywhere?

When $f \in L^1$, the answer is yes. The simplest proof of this fact proceeds by showing that the Fourier transform[1] of f

$$\hat{f}(\xi, \eta) = \iint f(x, y) e^{i(x\xi + y\eta)}\, dx\, dy$$

[1]We suppress the constant $\frac{1}{2\pi}$ as only the vanishing or nonvanishing of the Fourier transform will be of concern to us here.

vanishes identically. Indeed, if ℓ is a line through the origin, we may choose orthogonal coordinates in such a way that ℓ becomes the y-axis. Then, by Fubini's theorem,

$$\begin{aligned}\hat{f}(0,\eta) &= \iint e^{iy\eta} f(x,y)\,dx\,dy \\ &= \int e^{iy\eta}\left(\int f(x,y)\,dx\right)dy;\end{aligned} \tag{4.28}$$

and the inner integral vanishes for each fixed value of y. It follows that $\hat{f}$ vanishes on ℓ and hence (since ℓ was arbitrary) on each line through 0. Thus $\hat{f} = 0$, so by uniqueness $f = 0$. Since $\hat{f}$ is a continuous function for $f \in L^1$, this proof actually shows that it suffices for (4.27) to hold only for *almost* every line belonging to a *dense* set of directions.

Actually, the argument above can be worked backwards as well. If the left hand side of (4.28) is identically zero, the uniqueness theorem for the one-dimensional Fourier transform shows that $\int f(x,y)dx$ must vanish for almost every y. We conclude that the Fourier transform $\hat{f}$ vanishes on a line $\tilde{\ell}$ through the origin exactly when f satisfies (4.27) for almost all lines perpendicular to $\tilde{\ell}$. This observation will prove useful in the sequel. □

2. When the integrable function f vanishes off a bounded set, a much stronger result holds. In that case, the Fourier transform $\hat{f}$ is an entire function of ξ and η. Suppose $\hat{f}$ vanishes on an infinite collection of lines $\ell_1, \ell_2, \ldots$ through 0. Then, for each j, $\hat{f}$ is divisible by the linear factor L_j vanishing on ℓ_j. Writing $\boldsymbol{\xi} = (\xi, \eta)$, we then have $\hat{f}(\boldsymbol{\xi}) = O(|\boldsymbol{\xi}|^n)$ for each n, so the Maclaurin expansion of $\hat{f}$ into a series of homogeneous polynomials vanishes identically and $f = 0$. (Alternatively, assume without loss of generality that ℓ_j has the equation $\eta = \alpha_j\xi$ where $\alpha_j \to \alpha$. For fixed ξ, $g(z) = \hat{f}(\xi, z\xi)$ is an entire function with zeros at the α_j; thus $g(z) \equiv 0$, so $\hat{f}(\xi,\eta) = 0$ and $f \equiv 0$.) It follows that for functions of compact support, one need require only that (4.27) hold for almost all lines in each of an (arbitrary) infinite set of directions.

This conclusion obviously persists whenever the Fourier transform $\hat{f}$ is real analytic throughout $\mathbb{R}^2$. It is sufficient, for instance, that there exist positive constants K and c such that

$$|f(x,y)| \le Ke^{-c(|x|+|y|)}.$$

On the other hand, a function of compact support can satisfy (4.27) for all lines in a finite number of directions and still have an (almost) arbitrary shape. Indeed, given lines $\ell_1, \ell_2, \ldots, \ell_n$ through 0, we may choose a polynomial $P = P(\boldsymbol{\xi})$ which vanishes on the union of the ℓ_j; for instance, a product of linear factors will do. Let $D = (-i\partial/\partial x, -i\partial/\partial y)$. Then if g is an arbitrary smooth function of compact support and $f = P(D)g$, by a familiar fact from Fourier analysis we have

$$\hat{f}(\boldsymbol{\xi}) = (P(D)g)^\wedge(\boldsymbol{\xi}) = P(\boldsymbol{\xi})\hat{g}(\boldsymbol{\xi}),$$

which vanishes for $\boldsymbol{\xi} \in \ell_j$, $1 \le j \le n$. Thus (4.27) holds for all lines orthogonal to any one of the lines ℓ_j. □

3. In the general case, however, no improvement of the sort discussed in the preceding section is to be expected, even for functions in the Schwartz class $\mathscr{S}$ of

smooth, rapidly decreasing functions on $\mathbb{R}^2$. Indeed, given an arc α on the unit circle, take a disc D contained in the angle subtended at the origin by α. Choose a smooth function ϕ ($\phi \neq 0$) supported in D and let $f = \hat{\phi}$. Since $\phi \in \mathscr{S}$, $f \in \mathscr{S}$. Moreover, (4.27) holds for each line ℓ perpendicular to a direction *not* in α. Indeed, by Fourier inversion, we have

$$\hat{f}(\boldsymbol{\xi}) = \hat{\hat{\phi}}(\boldsymbol{\xi}) = (2\pi)^2\phi(-\boldsymbol{\xi}) = 0$$

on any line through 0 whose direction does not belong to α. The claim then follows from the remark at the end of §1.

Thus, for any open set of directions Θ, there exists a nonzero function in $\mathscr{S}$ satisfying (4.27) for all lines whose directions are *not* in Θ. It follows that the assumption that (4.27) holds for almost all lines in a dense set of directions cannot, in general, be relaxed, even for smooth functions which tend rapidly, with their derivatives, to zero. □

4. In the absence of measurability assumptions, the situation changes drastically: there exist *nonmeasurable* functions for which (4.27) holds. Indeed, Sierpiński [**Si**] has demonstrated the existence of a nonmeasurable set E with the property that each line intersects E in at most two points. The characteristic function χ_E then obviously satisfies (4.27) but is nonnull. If a function of bounded support with the same property is desired, it suffices to take $f = \chi_{E\cap D}$, where D is a sufficiently large disc about the origin.

The proof of Sierpiński's result is not difficult; if we assume the continuum hypothesis, it becomes [**O**, pp. 54–55] too simple to omit. Consider, then, the collection $\mathscr{F}$ of all closed subsets of $\mathbb{R}^2$ having positive planar measure. Since $\mathscr{F}$ has the power of the continuum, it may be well-ordered in such a way that each member F_α of $\mathscr{F}$ has only countably many predecessors, i.e., so that $\mathscr{F}$ has ordinal ω_1, the first uncountable ordinal. Choose $p_1 \in F_1$ and $p_2 \in F_2$, $p_2 \neq p_1$. Suppose that $\alpha < \omega_1$ and that $p_\beta \in F_\beta$ has been chosen for all $\beta < \alpha$. The point set $E_\alpha = \{p_\beta : \beta < \alpha\}$ is countable and therefore determines only countably many lines, whose union is then a set of planar measure zero. Since F_α has positive measure, we may choose $p_\alpha \in F_\alpha$ disjoint from any of these lines.

Set $E = \{p_\alpha : \alpha < \omega_1\}$. No three points of E are collinear, so it is clear that each line intersects E in at most two points. Suppose E is measurable. Then by Fubini's theorem, $m(E) = 0$. On the other hand, $\tilde{E}$ (the complement of E) is also measurable and

$$m(\tilde{E}) = \sup\{m(F) : F \subset \tilde{E},\ F \text{ closed}\}$$

by the regularity of Lebesgue measure. Since E meets every closed set of positive measure, $m(\tilde{E}) = 0$; and we have a contradiction. □

5. The example of the previous section is less than totally satisfying. For one thing, it is not, and cannot be, constructive. Indeed, there are models of set theory (in which the axiom of choice fails to hold) in which *every* set – and hence every function – is measurable [**So**]. More importantly, the example skirts the main point, focusing instead on the marginal issue of measurability. The real question is whether a reasonable (say, continuous) function can satisfy (4.27) for all lines without vanishing identically. It turns out that the answer is yes.

We shall, in fact, exhibit a nonzero *entire* function $g(z)$ which for every line ℓ in the plane satisfies

(i) $\int_\ell |g'(z)|ds < \infty$

and

(ii) $g(z) \to 0$ as $z \to \infty$ $(z \in \ell)$.

Then by (i), $f(z) = g'(z)$ is absolutely integrable on every line, while (ii) together with the fundamental theorem of calculus shows that (4.27) holds.

The original construction of such a function **[Z2]** utilized a deep theorem of Arakelian on tangential approximation by entire functions. Using the classical technique of "pole-pushing," David Armitage **[A1]** gave a beautifully simple construction, based on the following

LEMMA. *Let $z_1, z_2 \in \mathbb{C}$, $|z_1 - z_2| < 1$. Given a function h analytic on $\mathbb{C} \setminus \{z_1\}$ and $\varepsilon > 0$, there exists a function k analytic on $\mathbb{C} \setminus \{z_2\}$ such that*

$$|h(z) - k(z)| < \frac{\varepsilon}{(1+|z|)^2} \quad \textit{for} \quad |z - z_2| > 1. \tag{4.29}$$

PROOF. Expanding h in a Laurent series on $\{z : |z_1 - z_2| < |z - z_2|\}$, we have

$$h(z) = h_0(z) + \sum_{n=1}^{\infty} \frac{a_n}{(z - z_2)^n}, \tag{4.30}$$

where h_0 is entire and the series converges uniformly for $|z - z_2| \geq 1$. We claim that for sufficiently large m, the function

$$k(z) = h_0(z) + \sum_{n=1}^{m} \frac{a_n}{(z - z_2)^n} \tag{4.31}$$

satisfies (4.29). Indeed, from the continuity of the function $|z - z_2|/(1 + |z|)$ and the fact that $|z - z_2|/(1 + |z|) \to 1$ as $z \to \infty$, it follows that there exists a constant $C = C(z_2)$ such that

$$\frac{1}{|z - z_2|^2} < \frac{C}{(1+|z|)^2} \quad \text{for} \quad |z - z_2| > 1. \tag{4.32}$$

Since the series in (4.30) converges absolutely when $|z - z_2| > |z_1 - z_2|$ and $|z_1 - z_2| < 1$, $\sum\limits_{n=1}^{\infty} |a_n| < \infty$; so by choosing m sufficiently large, we can ensure that

$$\sum_{n=m+1}^{\infty} |a_n| < \varepsilon/C. \tag{4.33}$$

It then follows from (4.33) that for $|z - z_2| > 1$,

$$\left| \sum_{n=m+1}^{\infty} \frac{a_n}{(z - z_2)^n} \right| \leq \sum_{n=m+1}^{\infty} \frac{|a_n|}{|z - z_2|^n} \leq \sum_{n=m+1}^{\infty} \frac{|a_n|}{|z - z_2|^2} < \frac{\varepsilon}{C} \frac{1}{|z - z_2|^2},$$

which by (4.32) implies

$$\left| \sum_{n=m+1}^{\infty} \frac{a_n}{(z - z_2)^n} \right| < \frac{\varepsilon}{(1+|z|)^2} \quad \text{for} \quad |z - z_2| > 1,$$

as claimed. □

To construct g, choose points $\{z_n\}$ on the semiparabola $P = \{(x, x^2) : x \geq 0\}$ such that $z_0 = 0$, $|z_n - z_{n-1}| < 1$ for $n \geq 1$, and $z_n \to \infty$. Let $g_0(z) = 1/z^2$. Proceeding inductively, assume that g_k has been defined for $0 \leq k \leq n-1$. Applying the Lemma, we obtain a function g_n analytic on $\mathbb{C} \setminus \{z_n\}$ such that

$$(4.34) \qquad |g_n(z) - g_{n-1}(z)| < \frac{1}{2^n} \frac{1}{(1+|z|)^2} \quad \text{for} \quad |z - z_n| > 1.$$

The sequence $\{g_n\}$ converges uniformly on compacta to a limit function g, which is entire. Denote by P_a the set of all points in $\mathbb{C}$ whose distance from P is greater than $a > 0$. Then by (4.34),

$$|g(z) - g_0(z)| \leq \sum_{n=1}^{\infty} |g_n(z) - g_{n-1}(z)| < \frac{1}{(1+|z|)^2} < \frac{1}{|z|^2}$$

for $z \in P_1$. Hence $g \not\equiv 0$, and

$$(4.35) \qquad |g(z)| \leq \frac{2}{|z|^2}, \quad z \in P_1.$$

Cauchy's formula for derivatives then yields

$$(4.36) \qquad |g'(z)| \leq \frac{2}{|z|^2}, \quad z \in P_2.$$

Since $\ell \setminus P_2$ is a bounded set for each line ℓ, (i) and (ii) follow from (4.35) and (4.36). □

REMARKS. 1. The construction above yields considerably more than claimed: not only is f $(= g')$ absolutely integrable with integral 0 on every line ℓ, but so is each of its derivatives! Indeed, the analogue of (4.36) for higher derivatives shows that $|g^{(n)}(z)| \leq 2n!/|z|^2$ for $z \in P_2$, so that analogues of (i) and (ii) hold with g replaced by $g^{(n)}$, $n = 1, 2, \ldots$. Robert Burckel **[Bu]** has shown how the construction can be adjusted to yield an entire function which tends to zero, along with all its derivatives, as $z \to \infty$ along every (unbounded) algebraic curve.

2. Explicit examples of entire functions exhibiting the behavior discussed above are given in **[A2]**, which also contains references to the literature going back to Lindelöf and Mittag-Leffler.

3. For more recent developments, see **[BMR]** and **[Bo]**.

COMMENT. In its classical guise, the Radon transform associates to a function f defined in $\mathbb{R}^n$ the function

$$\hat{f}(\xi) = \int_{\xi} f(x) dm(x)$$

defined on the collection of $(n-1)$-dimensional affine subspaces ξ of $\mathbb{R}^n$ by integrating f over ξ against Lebesgue measure. Initiated in 1917 by J. Radon **[R]**, the theory has undergone a remarkable development during recent decades. The mapping properties of the Radon transform have been investigated, leading to results which parallel the more familiar theory of the Fourier transform. At the same time, analogues of the Radon transform have been defined on noneuclidean spaces, with interesting and appealing results. (That the noneuclidean theory raises challenges well beyond those of the classical theory is amply illustrated in **[LP]**.) The place to read about these developments is Helgason's authoritative (and recently updated) exposition **[H]**; see also **[St]** for an appetizing introduction to this material.

Readers interested in exploring the theory obtained when lines (or planes) are replaced by circles (or spheres) are directed to Fritz John's lovely little monograph [**J**] and, for more recent developments, [**Z1**].

Great interest has also focused on the subject from the point of view of "real life" applications, which range from radioastronomy to nuclear magnetic-resonance reconstructions. Most spectacular of all are the applications to medical radiology, viz., computed tomography. This last has been termed the most important development in diagnostic medicine since the discovery of x-rays, a judgment confirmed in part by the award of the 1979 Nobel Prize in Medicine and Physiology to A.M. Cormack and G.N. Hounsfield "for the development of computer assisted tomography."

Bibliography

[A1] D.H. Armitage, *A nonconstant continuous function on the plane whose integral on every line is zero*, Amer. Math. Monthly **101** (1994), 892-894.

[A2] D.H. Armitage, *Entire functions that tend to zero on every line*, Amer. Math. Monthly **114** (2007), 251-256.

[BMR] Claude Bélisle, Jean-Claude Massé and Thomas Ransford, *When is a probability measure determined by infinitely many projections?*, Ann. Probab. **25** (1997), 767-786.

[Bo] Jan Boman, *Unique continuation of microlocally analytic distributions and injectivity theorems for the ray transform*, Inverse Probl. Imaging **4** (2010), 619-630.

[Bu] R.B. Burckel, *Entire functions which vanish at infinity*, Amer. Math. Monthly **102** (1995), 916-918.

[H] Sigurdur Helgason, *Integral Geometry and Radon Transforms*, Springer, 2010.

[J] Fritz John, *Plane Waves and Spherical Means, Applied to Partial Differential Equations*, Interscience, 1955.

[LP] Peter D. Lax and Ralph S. Phillips, *A local Paley-Wiener theorem for the Radon transform of L_2 functions in a non-Euclidean setting*, Comm. Pure Appl. Math. **35** (1982), 531-554.

[O] John C. Oxtoby, *Measure and Category*, Springer, 1971.

[R] Johann Radon, *Über die Bestimmung von Fuktionen durch ihre Integralwerte längs gewisser Mannigfaltigkeiten*, Ber. Verh. Sächs Akad. Wiss. Leipzig Math.-Natur Kl. **69** (1917), 262-277.

[Si] Waclaw Sierpiński, *Sur un problème concernant les ensembles mesurables superficiellement*, Fund. Math. **1** (1920), 112-115.

[So] Robert M. Solovay, *A model of set-theory in which every set of reals is Lebesgue measurable*, Ann. of Math. (2) **92** (1970), 1-56.

[St] Robert S. Strichartz, *Radon inversion – Variations on a theme*, Amer. Math. Monthly **89** (1982), 161-175.

[Z1] Lawrence Zalcman, *Offbeat integral geometry*, Amer. Math. Monthly **87** (1980), 161-175.

[Z2] Lawrence Zalcman, *Uniqueness and nonuniqueness for the Radon transform*, Bull. London Math. Soc. **14** (1982), 241-245.

4.4. The Paley-Wiener Theorem

A number of results in harmonic analysis answer to the name of the Paley-Wiener Theorem. Typically, such results characterize the behavior of a function on the line in terms of the analyticity of its Fourier transform on some portion of the complex plane.

Suppose, for instance, that F is an integrable function supported on the positive half-line $\mathbb{R}_+ = [0, \infty)$. Then

$$f(w) = \int_0^\infty F(\xi) e^{i\xi w} d\xi \tag{4.37}$$

is a bounded analytic function in the upper half-plane $\{w = u + iv : v > 0\}$. When F vanishes on an interval $[0, \ell]$, more can be said. Indeed, in that case, setting $\xi = \sigma + \ell$ in (4.37), we can write

$$f(w) = \int_{\ell}^{\infty} F(\xi)e^{i\xi w}d\xi = e^{i\ell w}\int_0^{\infty} F(\sigma+\ell)e^{i\sigma w}d\sigma,$$

so that

$$|e^{-i\ell w}f(w)| \leq \int_0^{\infty} |F(\sigma+\ell)|d\sigma = \int_{\ell}^{\infty} |F(\xi)|d\xi \tag{4.38}$$

for $\operatorname{Im} w \geq 0$.

It turns out that the boundedness of the left hand side of (4.38) is actually equivalent to the vanishing of F on $[0, \ell]$. This is the

PALEY-WIENER THEOREM. *Let $F \in L^1(\mathbb{R}_+)$. Then $F(\xi) = 0$ for almost all $\xi \in [0, \ell]$ if and only if there exists a constant $A > 0$ such that*

$$|f(w)| \leq Ae^{-\ell v} \qquad \textit{for} \quad v = \operatorname{Im} w > 0. \tag{4.39}$$

PROOF. We have just seen that (4.38), which is equivalent to (4.39), holds if F vanishes on $[0, \ell]$. We prove the converse, that (4.39) implies that F is zero on $[0, \ell]$, by showing that for any smooth function G supported on $[0, \ell - d]$, $0 < d < \ell$,

$$(F, G) = \int_0^{\infty} F(\xi)\overline{G}(\xi)d\xi = 0. \tag{4.40}$$

To this end, let

$$g(u) = \int_0^{\ell-d} G(\xi)e^{i\xi u}d\xi,$$

so that

$$\overline{g}(u) = \int_0^{\ell-d} \overline{G}(\xi)e^{-i\xi u}d\xi. \tag{4.41}$$

We claim that

$$(F, G) = \frac{1}{2\pi}(f, g), \tag{4.42}$$

where

$$(f, g) = \int_{-\infty}^{\infty} f(u)\overline{g}(u)du. \tag{4.43}$$

Indeed, approximating F in L^1 norm by a sequence $\{F_n\}$ of functions in $L^1 \cap L^2$, we have by Parseval's formula

$$(F_n, G) = \frac{1}{2\pi}(f_n, g), \tag{4.44}$$

where

$$f_n(u) = \int_0^{\infty} F_n(\xi)e^{i\xi w}d\xi.$$

Since $f_n \to f$ uniformly on $\mathbb{R}$ and g (as the Fourier transform of a smooth function of compact support) belongs to the Schwartz class and hence is integrable, we may pass to the limit as $n \to \infty$ in (4.44) to obtain (4.42).

The formula (4.41) shows that $\overline{g}$ can be extended to the whole complex plane as an entire function

$$h(w) = \int_0^{\ell-d} \overline{G}(\xi)e^{-i\xi w}d\xi. \tag{4.45}$$

Since G is smooth and has compact support, we can integrate (4.45) by parts twice to obtain

$$|h(w)| \le \frac{Ce^{(\ell-d)v}}{1+|w|^2} \qquad v = \operatorname{Im} w > 0, \tag{4.46}$$

for some constant $C > 0$. Now f is analytic and bounded in the upper half-plane, so by Cauchy's Theorem and the estimate (4.46), we can shift the line of integration in (4.43) from the real axis to the line $\operatorname{Im} w = v > 0$:

$$(f,g) = \int_{-\infty}^{\infty} f(u)h(u)du = \int_{-\infty}^{\infty} f(u+iv)h(u+iv)du. \tag{4.47}$$

The bounds in (4.39) and (4.46) give

$$|f(w)h(w)| \le \frac{AC}{1+|w|^2}e^{-dv},$$

so the right hand side of (4.47) tends to 0 as $v \to \infty$. Since the left hand side is independent of v, it must vanish. Hence, by (4.42), $(F,G) = 0$ for all smooth functions supported on some compact subinterval of $[0,\ell)$. It follows that F is zero on $[0,\ell]$, as claimed. $\square$

COMMENT. As mentioned above, there are a number of results which go by the name of Paley-Wiener Theorem. Perhaps the best-known of these is the following.

THEOREM. [**PW**, pp. 11-12] *Let $A > 0$. Then*

$$f(u) = \int_{-A}^{A} F(\xi)e^{i\xi u}d\xi \tag{4.48}$$

for some $F \in L^2[-A,A]$ if and only if $f \in L^2(\mathbb{R})$ and f can be extended to the complex plane as an entire function of exponential type at most A.

A detailed, self-contained proof of this result is in [**Ch**, pp. 116-122]. The requirement that f extend to be of exponential type at most A means that for each $\varepsilon > 0$, there exists $C_\varepsilon > 0$ such that

$$|f(w)| \le C_\varepsilon e^{(A+\varepsilon)|w|}, \qquad w \in \mathbb{C}.$$

In point of fact, if (4.48) holds, then

$$f(w) = o(e^{A|w|}) \qquad \text{as} \quad w \to \infty.$$

Bibliography

[Ch] K. Chandrasekharan, *Classical Fourier Transforms*, Springer-Verlag, 1989.
[PW] Raymond E.A.C. Paley and Norbert Wiener, *Fourier Transforms in the Complex Plane*, American Mathematical Society, 1934.

4.5. The Titchmarsh Convolution Theorem

We consider integrable functions on the positive half-line $\mathbb{R}_+ = [0, \infty)$. Denote the lower end of the support of such a function F by

$$\ell_F = \max\{\eta : F(\xi) = 0 \quad \text{for a.a.} \quad \xi < \eta\}.$$

A celebrated result of Titchmarsh **[T]** describes the behavior of ℓ_F under the operation of convolution.

THEOREM. *Let $A, B \in L^1(\mathbb{R}_+)$ and denote by*

$$(A * B)(\xi) = \int_0^\xi A(\eta)B(\xi - \eta)d\eta \tag{4.49}$$

their convolution. Then

$$\ell_{A*B} = \ell_A + \ell_B. \tag{4.50}$$

PROOF. For $\xi < \ell_A + \ell_B$, the integrand on the right of (4.49) is zero, since a.e. at least one of the factors is zero. Therefore, the integral is zero, i.e., $(A * B)(\xi) = 0$ for $\xi < \ell_A + \ell_B$, so that

$$\ell_A + \ell_B \leq \ell_{A*B}. \tag{4.51}$$

It remains to prove that equality holds in (4.51).

To this end, let us recall that, according to the Paley-Wiener Theorem,

$$\ell_F = \max\{\ell : |f(w)e^{-i\ell w}| \leq C \qquad \text{for some} \quad C > 0\},$$

where

$$f(w) = \int_0^\infty F(\xi)e^{i\xi w}d\xi.$$

Equivalently, in the language of division in the algebra $\mathcal{B}$ of bounded analytic functions in the upper half-plane, ℓ_F is the highest power of e^{iw} that divides the Fourier transform of F in $\mathcal{B}$.

Applying this to the situation at hand, let us denote the Fourier transforms of A and B by $a(w)$ and $b(w)$, respectively; these are elements of $\mathcal{B}$. The Fourier transform of the convolution $A * B$ is then $\sqrt{2\pi}ab$. Thus, (4.50) can be restated by saying that if ℓ_A and ℓ_B denote the highest powers of e^{iw} that divide the functions $a(w)$ and $b(w)$, respectively, then the highest power that divides their product ab is $\ell_A + \ell_B$.

To prove this, we factor a and b as $a = e^{i\ell_A w}c$ and $b = e^{i\ell_B w}d$, where c and d belong to $\mathcal{B}$. The functions c and d are relatively prime to e^{iw}. Indeed, according to Theorem 3.13, any divisor of e^{iw} has the form e^{ikw}, $k > 0$; on the other hand, neither c nor d has divisors of that form, for then a or b would be divisible by a higher power of e^{iw} than stipulated. It now follows from Theorem 3.12 that the product cd is relatively prime to e^{iw}. This shows that $ab = e^{i(\ell_A+\ell_B)w}cd$ is not divisible by a power of e^{iw} greater than $\ell_A + \ell_B$ and completes the proof of the theorem. □

COMMENT. It is a curious fact that although Titchmarsh's Convolution Theorem is a real-variable result, Titchmarsh's original proof **[T]** used complex variable theory, as did all subsequent proofs for the next quarter century. Only in 1952 did Mikusiński and Ryll-Nardzewski discover a proof avoiding complex analysis. Fairly simply elementary proofs are now available in **[M**, Chapter XV] and

[D]. The proof given above, which is taken from [L], shows that the approach via complex variables is not unnatural.

Bibliography

[D] Raouf Doss, *An elementary proof of Titchmarsh's Convolution Theorem*, Proc. Amer. Math. Soc. **104** (1988), 181-184.

[L] Peter D. Lax, *Translation invariant spaces*, Acta Math. **101** (1959), 163-178.

[M] Jan Mikusiński, *The Bochner Integral*, Birkhäuser Verlag, 1978.

[T] E.C. Titchmarsh, *The zeros of certain integral functions*, Proc. London Math. Soc. (2) **25** (1926), 283-302.

4.6. Hardy's Theorem

Recall that the Fourier transform of $f \in L^1(\mathbb{R})$ is

$$(\mathcal{F}f)(u) = \frac{1}{\sqrt{2\pi}} \int f(x) e^{ixu} dx = \hat{f}(u).$$

In particular, $(e^{-\alpha x^2})^\wedge = (1/\sqrt{2\alpha}) e^{-u^2/4\alpha}$, so that $(e^{-x^2/2})^\wedge = e^{-u^2/2}$. More generally, if H_n is the nth Hermite polynomial and

$$\text{(a)} \qquad \varphi_n(x) \doteq e^{-x^2/2} H_n(x) = (-1)^n e^{x^2/2} \left(\frac{d}{dx}\right)^n e^{-x^2},$$

then $\hat{\varphi}_n(u) = (i)^n \varphi_n(u)$.

According to a general principle of harmonic analysis (attributed by G.H. Hardy to Norbert Wiener), a nonzero function and its Fourier transform cannot both be very small. One instance of this phenomenon is the celebrated Uncertainty Principle [HJ]. Another is the following beautiful theorem due to Hardy [Ha].

THEOREM. *Let $f \in L^1(\mathbb{R})$ and suppose that there exist positive constants C', C'', α and β such that*

$$\text{(b)} \qquad |f(x)| \le C' e^{-\alpha x^2} \quad \textit{and} \quad |\hat{f}(u)| \le C'' e^{-\beta u^2}$$

for all $x, u \in \mathbb{R}$. Then

(1) *if $\alpha\beta = 1/4$, $f(x) = Ae^{-\alpha x^2}$ and $\hat{f}(u) = (A/\sqrt{2\alpha})e^{-u^2/4\alpha}$ for some A;*
(2) *if $\alpha\beta > 1/4$, $f = \hat{f} = 0$;*
(3) *for $\alpha\beta < 1/4$, there exist infinitely many such functions f.*

The proof depends on an application of the Phragmén-Lindelöf Principle (Appendix C) and Liouville's Theorem. It is convenient to state this part of the argument as a separate result.

LEMMA. *Let g be an entire function. Suppose there exist positive numbers C and a such that*

(i) *$|g(w)| \le Ce^{a|w|}$ for all $w \in \mathbb{C}$;*
(ii) *$|g(u)| \le Ce^{-au}$ for $u > 0$.*

Then $g(w) = Ae^{-aw}$ $(w \in \mathbb{C})$ for some constant A.

PROOF. Taking $\delta > 0$ small and applying the Phragmén-Lindelöf Theorem to the function

$$F_\delta(w) = g(w) \exp\left[\left(a + ia\tan\frac{\delta}{2}\right) w\right]$$

on the angle $D_\delta = \{w \in \mathbb{C} : 0 < \arg w < \pi - \delta\}$ (as we may, since F_δ clearly has order at most 1), we obtain

$$\sup_{D_\delta} |F_\delta(w)| \le \max\left\{\sup_{u>0} |F_\delta(u)|, \sup_{r>0} |F_\delta(re^{i(\pi-\delta)})|\right\}.$$

Now by (ii),

$$|F_\delta(u)| = |g(u)|e^{au} \le Ce^{-au}e^{au} = C$$

for $u > 0$. On the other hand, for $w = re^{i(\pi-\delta)} = -r\cos\delta + ir\sin\delta$, we have

$$\operatorname{Re}\left(a + ia\tan\frac{\delta}{2}\right) w = -ar\left(\cos\delta + \tan\frac{\delta}{2}\sin\delta\right) = -ar.$$

For such w, (i) gives

$$\begin{aligned}|F_\delta(w)| &\le |g(w)| \left|\exp\left(a + ia\tan\frac{\delta}{2}\right) w\right| \\ &\le Ce^{ar}e^{-ar} = C.\end{aligned}$$

It follows that for each $0 < \delta < \pi$, $|F_\delta(w)| \le C$ on D_δ. Thus

$$|g(w)e^{aw}| = \lim_{\delta \to 0} |F_\delta(w)| \le C$$

on the upper half plane $v \ge 0$. A similar argument shows that the same estimate holds on the lower half plane. Thus $|g(w)e^{aw}| \le C$ throughout $\mathbb{C}$, so by Liouville's Theorem, $g(w)e^{aw}$ is constant. □

We now turn to the

PROOF OF HARDY'S THEOREM. A simple change of scale in the variables shows that we may assume $\alpha = \beta$. Let us first prove (1), which is the key result. So suppose $\alpha = \beta = 1/2$. Then

$$\hat{f}(w) = \frac{1}{\sqrt{2\pi}} \int f(x)e^{ixw}dx$$

is an entire function of $w = u + iv$, and we have

$$\begin{aligned}|\hat{f}(w)| &\le \frac{1}{\sqrt{2\pi}} \int |f(x)|e^{-xv}dx \\ &\le \frac{1}{\sqrt{2\pi}} \int C'e^{-x^2/2}e^{-xv}dx \\ &= \frac{C'}{\sqrt{2\pi}} \int e^{-(x+v)^2/2}dx \cdot e^{v^2/2} \\ &= C'e^{v^2/2}.\end{aligned}$$

Suppose f is even. Then $\hat{f}$ is also even, and it follows that $g(w) = \hat{f}(\sqrt{w})$ is an entire function. Since $|\hat{f}(w)| \le C'e^{v^2/2}$,

$$|g(w)| \le C' \exp\left[\frac{1}{2}(\operatorname{Im}\sqrt{w})^2\right] \le C'e^{|w|/2}.$$

Moreover, since $|\hat{f}(u)| \leq C'' e^{-u^2/2}$ by assumption, $|g(u)| \leq C'' e^{-u/2}$ for $u > 0$. Replacing C' and C'' by $C = \max(C', C'')$ and applying the Lemma, we obtain $g(w) = Ae^{-w/2}$, so that $\hat{f}(w) = g(w^2) = Ae^{-w^2/2}$ and hence also $f(x) = Ae^{-x^2/2}$. This completes the proof of (1) in case f is even.

If f is odd, then $\hat{f}$ is also odd; so $\hat{f}(0) = 0$, and we can apply the previous proof to the even entire function $\hat{f}(w)/w$ to obtain $\hat{f}(w)/w = Ae^{-w^2/2}$. Since $|\hat{f}(u)| \leq C'' e^{-u^2/2}$ for u real, we must have $A = 0$, so $\hat{f} = 0 = f$.

In general, we decompose $f = f_e + f_o$ into its even and odd parts, each of which then satisfies the hypotheses of the theorem, and apply the arguments above to f_e and f_o separately to see that $f_o = \hat{f}_o = 0$ and $f(x) = f_e(x) = Ae^{-x^2/2}$.

It remains to prove (2) and (3). Suppose first that $\alpha\beta > 1/4$, as assumed in (2). Normalizing by $\alpha = \beta$, we have $\alpha = \beta > 1/2$, so the assumptions $|f(x)| \leq C' e^{-\alpha x^2}$ and $|\hat{f}(u)| \leq C'' e^{-\beta u^2}$ imply that $|f(x)| \leq C' e^{-x^2/2}$ and $|\hat{f}(u)| \leq C'' e^{-u^2/2}$ for all $x, u \in \mathbb{R}$. By part (1), $f(x) = Ae^{-x^2/2}$. But this is consistent with (b) only if $A = 0$.

Finally, suppose $\alpha\beta < 1/4$. Normalizing again, we may assume $\alpha = \beta < 1/2$. Let φ_n be the Hermite function of (a). Then there exists a positive constant C (depending on n) such that

$$|\varphi_n(x)| \leq C(1 + |x|^n)e^{-x^2/2} \qquad \text{and} \qquad |\hat{\varphi}_n(u)| \leq C(1 + |u|^n)e^{-u^2/2}$$

for all $x, u \in \mathbb{R}$. It follows that for each $\alpha < 1/2$, there exists $C' = C'(\alpha, n)$ such that

$$|\varphi_n(x)| \leq C' e^{-\alpha x^2} \qquad \text{and} \qquad |\hat{\varphi}_n(u)| \leq C' e^{-\alpha u^2}$$

for all $x, u \in \mathbb{R}$. □

COMMENTS. 1. Hardy also showed that if

$$|f(x)| = O(|x|^m e^{-x^2/2}) \qquad \text{and} \qquad |\hat{f}(u)| = O(|u|^m e^{-u^2/2})$$

for large x and u and some positive integer m, then both f and $\hat{f}$ are finite linear combinations of Hermite functions.

2. G.W. Morgan [**M**] proved an extension of part (2) of Hardy's Theorem. He showed that if $p > 2$, $1/p + 1/q = 1$ and $A > 0$, there exists $A' > 0$ (depending on A and p in a specific manner) such that for each $\varepsilon > 0$, the conditions

$$|f(x)| \leq C' e^{-Ax^p} \qquad \text{and} \qquad |\hat{f}(u)| \leq C'' e^{-(A'+\varepsilon)u^q}$$

imply that $f = \hat{f} = 0$.

3. Another result of this sort, less well-known than it should be, is the following striking theorem of Beurling [**B**, p. 372].

THEOREM. *Let $f \in L^1(\mathbb{R})$ and suppose*

$$\iint_{\mathbb{R}^2} |f(x)\hat{f}(u)| e^{|xu|} dx du < \infty.$$

Then $f(x) = 0$ a.e. on $\mathbb{R}$.

For the proof, which again uses the Phragmén-Lindelöf Principle, see [**Hö**] or [**L**, pp. 197-199].

Bibliography

[B] Arne Beurling, *The Collected Works of Arne Beurling*, Vol. 2, Birkhäuser Boston, 1989.

[Ha] G.H. Hardy, *A theorem concerning Fourier transforms*, J. London Math. Soc. **8** (1933), 227–231.

[HJ] Victor Havin and Burglind Jöricke, *The Uncertainty Principle in Harmonic Analysis*, Springer, 1994.

[Hö] Lars Hörmander, *A uniqueness theorem of Beurling for Fourier transform pairs*, Ark. Mat. **29** (1991), 237–240.

[L] B.Ya. Levin, *Lectures on Entire Functions*, American Mathematical Society, 1996.

[M] G.W. Morgan, *A note on Fourier transforms*, J. London Math. Soc. **9** (1934), 187–192.

CHAPTER 5

Banach Algebras: The Gleason-Kahane-Żelazko Theorem

Let A be a commutative Banach algebra with unit, and let M be a maximal ideal of A. Then M is a closed subspace, and the quotient Banach algebra A/M is a field which, by the Gelfand-Mazur Theorem, is isometrically isomorphic to $\mathbb{C}$. Thus the quotient map $\varphi : A \to A/M \cong \mathbb{C}$ is a complex homomorphism (multiplicative linear functional) of A. It follows that M is a (closed) linear subspace of codimension 1 in A which contains no invertible elements (since no proper ideal can contain invertible elements). The remarkable fact that this property characterizes maximal ideals was discovered by Gleason [**G**] and, independently, Kahane and Żelazko [**KŻ**].

THEOREM. *Let A be a commutative Banach algebra with unit. A linear subspace M of codimension 1 in A is a maximal ideal of A if and only if it contains no invertible elements.*

PROOF. We have already noted that a maximal ideal has the properties stated in the theorem. Suppose then that the linear subspace M has codimension 1 in A and contains no invertible elements. Then it contains no elements near the identity e, so its closure $\overline{M}$ is a proper subspace. Since M has codimension 1, $M = \overline{M}$ so M is closed. Let φ be the continuous linear functional on A such that $M = \ker \varphi$ and $\varphi(e) = 1$. We prove that

$$\varphi(xy) = \varphi(x)\varphi(y) \qquad x, y \in A, \tag{5.1}$$

from which it follows immediately that M is an ideal.

To this end, fix $x \in A$ and consider the analytic function f defined by

$$f(\lambda) = \varphi(e^{\lambda x}) = \sum_{n=0}^{\infty} \frac{\varphi(x^n)}{n!}\lambda^n. \tag{5.2}$$

Since $|\varphi(x^n)| \leq \|\varphi\| \, \|x\|^n$, f is entire and satisfies

$$|f(\lambda)| \leq \|\varphi\| e^{\|x\| \, |\lambda|}.$$

Moreover, since $\exp \lambda x$ is invertible in A, $f(\lambda) \neq 0$ for all $\lambda \in \mathbb{C}$; and $f(0) = \varphi(e) = 1$. Therefore, by the Corollary in Appendix B,

$$f(\lambda) = e^{\alpha\lambda} = \sum_{n=0}^{\infty} \frac{\alpha^n}{n!}\lambda^n \tag{5.3}$$

for some $\alpha \in \mathbb{C}$. Comparing (5.2) and (5.3), we have $\varphi(x^n) = \alpha^n$ for all n. In particular,

$$\varphi(x^2) = \varphi(x)^2. \tag{5.4}$$

Since this holds for each $x \in A$, it follows that for all $x, y \in A$,

$$\varphi((x+y)^2) = (\varphi(x) + \varphi(y))^2 \tag{5.5}$$

which, after simplification, reduces to (5.1), as required. □

REMARKS. 1. The proof given above actually shows a bit more than was claimed, in that it suffices to assume only that M contains no elements of the form e^x for $x \in A$.

2. If one assumes that A is an algebra over $\mathbb{R}$ rather than $\mathbb{C}$, the theorem is no longer true. A simple counterexample is obtained by taking the algebra $C_{\mathbb{R}}[0,1]$ of continuous real-valued functions on the unit interval and choosing

$$\varphi(f) = \int_0^1 f(t)dt.$$

Obviously, if $f \in C_{\mathbb{R}}[0,1]$ does not vanish, it has a single sign and thus $\varphi(f) \neq 0$. Hence $M = \ker\varphi$ contains no invertible elements. But $\varphi(f^2) > 0$ for *any* $f \not\equiv 0$, so φ is clearly not multiplicative.

COMMENTS. 1. This is not the end of the story. Żelazko noticed [**Ż**] that a further reasoning yields the fact that φ *must satisfy* (5.1) *even when A is not commutative!* The following simple argument is due to Rudin [**R**, pp. 251-252].

Note that the commutativity of A was used only in passing from (5.5) to (5.1); in the general case, we have (as was noted already by Gleason)

$$\varphi(xy + yx) = 2\varphi(x)\varphi(y) \qquad x, y \in A, \tag{5.6}$$

i.e., φ is a Jordan homomorphism. To show that (5.6) implies (5.1), suppose first that $\varphi(x) = 0$. Then it follows from (5.6) that

$$\varphi(xy + yx) = 0; \tag{5.7}$$

hence by (5.4),

$$\varphi((xy + yx)^2) = 0. \tag{5.8}$$

Writing

$$(xy - yx)^2 = 2(x(yxy) + (yxy)x) - (xy + yx)^2,$$

we have by (5.8) and (5.6),

$$\varphi((xy - yx)^2) = 2\varphi(x(yxy) + (yxy)x) = 4\varphi(x)\varphi(yxy) = 0;$$

so by (5.4),

$$\varphi(xy - yx) = 0. \tag{5.9}$$

Adding (5.7) and (5.9) then gives

$$\varphi(xy) = 0 \qquad \text{if} \qquad \varphi(x) = 0.$$

To complete the proof, let $x, y \in A$ be arbitrary. Since $\varphi(x - \varphi(x)e) = 0$,

$$0 = \varphi((x - \varphi(x)e)y) = \varphi(xy - \varphi(x)y) = \varphi(xy) - \varphi(x)\varphi(y),$$

which proves (5.1).

2. For a short and completely elementary proof of the GKŻ Theorem, see the paper by Roitman and Sternfeld [**RS**, pp. 112-113]. A useful list of related literature (until 1994) is in [**P**, p. 242]; cf. also [**R**, pp. 406-407].

3. Other attractive applications of complex function theory to Banach algebras appear in [**L**, §§6.2, 6.3, 28.3].

Bibliography

[G] Andrew M. Gleason, *A characterization of maximal ideals*, J. Anal. Math. **19** (1967), 171-172.

[KŻ] J.-P. Kahane and W. Żelazko, *A characterization of maximal ideals in commutative Banach algebras*, Studia Math. **29** (1968), 339-343.

[L] B.Ya. Levin, *Lectures on Entire Functions*, American Mathematical Society, 1996.

[P] Theodore W. Palmer, *Banach Algebras and the General Theory of *-Algebras, Vol. 1, Algebras and Banach Algebras*, Cambridge University Press, 1994.

[R] Walter Rudin, *Functional Analysis*, second edition, McGraw-Hill, 1991.

[RS] M. Roitman and Y. Sternfeld, *When is a linear functional multiplicative?*, Trans. Amer. Math. Soc. **267** (1981), 111-124.

[Ż] W. Żelazko, *A characterization of multiplicative linear functionals in complex Banach algebras*, Studia Math. **30** (1968), 83-85.

CHAPTER 6

Complex Dynamics: The Fatou-Julia-Baker Theorem

Complex dynamics is the study of the iteration of analytic functions. For rational functions on the Riemann sphere $\hat{\mathbb{C}}$, the main lines of the theory were laid down by Pierre Fatou and Gaston Julia, working independently, in the last years of the second decade of the twentieth century. A bit later, Fatou also initiated the study of the iteration of transcendental entire functions in the plane, a line of investigation advanced notably (after a hiatus of 40 years) by I.N. Baker. More recently, under the impetus provided by the availability of computer graphics, the subject has entered a period of renewed activity, which continues to this day. Here we show how certain developments in the theory of normal families (elaborated in Appendix D) lead to a much simplified proof of one of the central results of the theory.

Let f be a rational function of degree $d \geq 2$ or a nonlinear entire function. We consider the family $\mathscr{F}$ of iterates $\{f^n : n \in \mathbb{N}\}$, where $f^1 = f$ and $f^n = f \circ f^{n-1}$. A point z is called *periodic* if $f^n(z) = z$ for some $n \in \mathbb{N}$; it is *repelling* if, in addition, $|(f^n)'(z)| > 1$. (When $z = \infty$, this last definition must be modified; cf. [**St**, pp. 25-26].) The *Fatou set* $\mathcal{F}(f)$ is the largest open set (in $\hat{\mathbb{C}}$ if f is rational, otherwise in $\mathbb{C}$) on which $\mathscr{F}$ is normal; its complement $\mathcal{J} = \mathcal{J}(f)$ is the *Julia set*. It is well-known, and easy to prove, that $\mathcal{J}$ and $\mathcal{F}$ are completely invariant, i.e., that $z \in \mathcal{J}$ if and only if $f(z) \in \mathcal{J}$ and similarly for $\mathcal{F}$ and that $\mathcal{J}(f^m) = \mathcal{J}(f)$ for each $m \in \mathbb{N}$ [**CG**, p. 56], [**St**, pp. 28-29]. Moreover, $\mathcal{J}(f)$ contains no isolated points [**CG**, p. 57], [**Bm**, pp. 554]; cf. [**Bw1**, pp. 159-160].

Our aim is to prove the following fundamental result, due (independently) to Fatou [**F1**] and Julia [**J**] for rational functions of degree $d \geq 2$ and to Baker [**Bk**] for transcendental entire functions.

THEOREM 6.1. *$\mathcal{J}(f)$ is the closure of the set of repelling periodic points of f.*

For the proof, we require some notation. The (forward) orbit of a point z is the set

$$O^+(z) = \{f^n(z) : n \in \mathbb{N}\}.$$

The backward orbit of z is the set of preimages of z under the iterates of f:

$$O^-(z) = \bigcup_{n=1}^{\infty} f^{-n}(\{z\}).$$

In general, for $S \subset \mathbb{C}$,

$$O^+(S) = \bigcup_{z \in S} O^+(z) \quad \text{and} \quad O^-(S) = \bigcup_{z \in S} O^-(z).$$

We have the following simple result.

THEOREM 6.2. *Let D be an open set such that $D \cap \mathcal{J} \neq \emptyset$. Then $\mathcal{J} \cap O^-(D)$ is a relatively open, dense subset of $\mathcal{J}$.*

PROOF. Since f is continuous, $O^-(D)$ is an open set. To see that $O^-(D) \cap \mathcal{J}$ is dense in $\mathcal{J}$, note first that since $\mathcal{J}$ contains no isolated points and $D \cap \mathcal{J} \neq \emptyset$, $D \cap \mathcal{J}$ must contain infinitely many points. If $O^-(D) \cap \mathcal{J}$ fails to be dense in $\mathcal{J}$, there exists an open set U such that $U \cap \mathcal{J} \neq \emptyset$ but $O^-(D) \cap \mathcal{J}$ and $U \cap \mathcal{J}$ are disjoint. Since $f^m(z) \in \mathcal{J}$ whenever $z \in \mathcal{J}$, this means that if $z \in U \cap \mathcal{J}$, then $f^m(z) \notin D$ for $m \in \mathbb{N}$, i.e., that $O^+(U \cap \mathcal{J})$ does not intersect D. By complete invariance, $O^+(U \cap \mathcal{F})$ is disjoint from $\mathcal{J}$. Thus $O^+(U) = O^+(U \cap \mathcal{J}) \cup O^+(U \cap \mathcal{F})$ does not intersect $D \cap \mathcal{J}$ and hence omits at least 3 (actually, infinitely many) values. It then follows from Montel's Theorem (Appendix D) that $\mathscr{F}$ is normal throughout U, which contradicts $U \cap \mathcal{J} \neq \emptyset$. □

We can now prove the Fatou-Julia-Baker Theorem.

PROOF OF THEOREM 6.1. Let z_0 be a repelling periodic point of f with period $p \in \mathbb{N}$. Without loss of generality, we may assume $z_0 \in \mathbb{C}$. Then $f^p(z_0) = z_0$ and $|(f^p)'(z_0)| > 1$, so by the chain rule,

$$\lim_{n \to \infty} |(f^{np})'(z_0)| = \infty.$$

It follows that no subsequence of $\{f^{np}\}$ is uniformly convergent on a neighborhood of z_0, i.e., $z_0 \in \mathcal{J}(f^p) = \mathcal{J}$. Since $\mathcal{J}$ is closed, the closure of the set of repelling periodic points of f is contained in $\mathcal{J}$.

To prove the opposite inclusion, consider the set $\mathscr{M}$ of points in $\mathcal{J}$ which are recurrent but not periodic, i.e., the set of all $z \in \mathcal{J}$ such that z belongs to the closure of $O^+(z) \setminus \{z\}$. We claim that $\mathscr{M}$ is dense in $\mathcal{J}$. Since $\mathcal{J}$ contains no isolated points, it suffices to show that the set $\{z \in \mathcal{J} : O^+(z) \text{ is dense in } \mathcal{J}\}$ is dense in $\mathcal{J}$. This turns out to be an easy consequence of the Baire Category Theorem. Indeed, for each $n \in \mathbb{N}$, we can cover $\mathcal{J}$ by at most countably many disks of radius $1/n$ to obtain altogether countably many disks D_j, each of which has nonempty intersection with $\mathcal{J}$. By Theorem 6.2, $Q_j = \mathcal{J} \cap O^-(D_j)$ is a relatively open, dense subset of $\mathcal{J}$. Applying Baire's Theorem to the complete metric space $\mathcal{J}$, we conclude that $Q = \bigcap Q_j$ is also dense in $\mathcal{J}$. Now suppose $q \in Q$. Then $O^+(q) \cap D_j \neq \emptyset$ for each j and hence $O^+(q)$ is dense in $\mathcal{J}$.

It remains to prove that $\mathscr{M}$ is contained in the closure of the set of repelling periodic points of f. To this end, suppose that $z_0 \in \mathscr{M}$ and let U be a neighborhood of z_0; we shall show that U contains a repelling fixed point of f. It is no loss of generality to assume that $z_0 \in \mathbb{C}$. Since $\{f^n\}$ is not normal on U, it follows from Zalcman's Lemma (Appendix D) that there exist points $z_k \to z_0$, numbers $\rho_k \to 0^+$ and an increasing sequence $\{n_k\}$ of positive integers such that

$$f^{n_k}(z_k + \rho_k \zeta) \to g(\zeta), \tag{6.1}$$

where g is a nonconstant meromorphic function and the convergence is uniform on compact subsets of the plane disjoint from the poles of g. By the definition of $\mathscr{M}$ and Picard's Theorem, there exists $m \in \mathbb{N}$ such that $f^m(z_0) \in U \cap g(\mathbb{C})$. Let $w_0 \in g^{-1}(f^m(z_0))$. Then there exists a neighborhood V of w_0 such that $g(V) \subset U$ and $g'(\zeta) \neq 0$ for all $\zeta \in V \setminus \{w_0\}$. Since $z_0 \in \mathscr{M}$, $f^m(z_0) \in \mathscr{M}$ too. Therefore, there exists $\ell \in \mathbb{N}$ and $\zeta_0 \in V \setminus \{w_0\}$ such that $g(\zeta_0) = f^\ell(z_0)$. Thus ζ_0 is an isolated

zero of the function

$$h(\zeta) = g(\zeta) - f^{\ell}(z_0) = \lim_{k\to\infty} \left[f^{n_k}(z_k + \rho_k\zeta) - f^{\ell}(z_k + \rho_k\zeta)\right].$$

By Hurwitz's Theorem, there exist points $\zeta_k \to \zeta_0$ such that

$$f^{n_k}(z_k + \rho_k\zeta_k) = f^{\ell}(z_k + \rho_k\zeta_k)$$

for all k sufficiently large. Thus $p_k = f^{\ell}(z_k + \rho_k\zeta_k)$ is a fixed point of $f^{n_k-\ell}$, hence a periodic point of f, for all large k. Differentiating (6.1) and using the fact that $\zeta_k \to \zeta_0$, we have

$$\begin{aligned} g'(\zeta_0) &= \lim_{k\to\infty} \frac{d}{d\zeta}\left[f^{n_k}(z_k + \rho_k\zeta)\right]\big|_{\zeta=\zeta_k} \\ &= \lim_{k\to\infty} \frac{d}{d\zeta}\left[f^{n_k-\ell}(f^{\ell}(z_k + \rho_k\zeta))\right]\big|_{\zeta=\zeta_k} \\ &= \lim_{k\to\infty} (f^{n_k-\ell})'(p_k)\cdot(f^{\ell})'(z_k+\rho_k\zeta_k)\cdot\rho_k. \end{aligned} \tag{6.2}$$

Now

$$\lim_{k\to\infty} (f^{\ell})'(z_k + \rho_k\zeta_k)\cdot\rho_k = (f^{\ell})'(z_0)\cdot 0 = 0. \tag{6.3}$$

On the other hand, since $\zeta_0 \in V \setminus \{w_0\}$, $g'(\zeta_0) \neq 0$; so it follows from (6.2) and (6.3) that

$$\lim_{k\to\infty} (f^{n_k-\ell})'(p_k) = \infty.$$

Thus, all but at most finitely many of the periodic points p_k are repelling. We complete the proof by noting that

$$\lim_{k\to\infty} p_k = \lim_{k\to\infty} f^{\ell}(z_k + \rho_k\zeta_k) = f^{\ell}(z_0) = g(\zeta_0) \in U.$$

□

COMMENT. Baker's original proof of Theorem 6.1 invokes the Ahlfors Five Islands Theorem, considered by many to be one of the deepest results in complex function theory. It was another thirty years before simpler proofs were found, first by Schwick [**Sk**], and then by Bargmann [**Bm**], whose argument we have followed above, and by Berteloot and Duval [**BD**]. While these proofs differ in significant detail, they all make essential use of Zalcman's Lemma. Inspired by some of this work, Bergweiler was led to a new (and much simpler) proof of Ahlfors' result [**Bw2**], which (again) hinges on Zalcman's Lemma.

Bibliography

[Bk] I.N. Baker, *Repulsive fixpoints of entire functions*, Math. Z. **104** (1968), 252-256.

[Bm] Detlef Bargmann, *Simple proofs of some fundamental properties of the Julia set*, Ergodic Theory Dynam. Systems **19** (1999), 553-558.

[Bw1] Walter Bergweiler, *Iteration of meromorphic functions*, Bull. Amer. Math. Soc. (N.S.) **29** (1993), 151-188.

[Bw2] Walter Bergweiler, *A new proof of the Ahlfors Five Islands Theorem*, J. Anal. Math. **76** (1998), 337-347.

[BD] François Berteloot and Julien Duval, *Une démonstration directe de la densité des cycles repulsifs dans l'ensemble de Julia*, Complex Analysis and Geometry, Birkhäuser, 2000, pp. 221-222.

[CG] Lennart Carleson and Theodore W. Gamelin, *Complex Dynamics*, Springer-Verlag, 1993.

[F1] P. Fatou, *Sur les équations fonctionelles*, Bull. Soc. Math. France **47** (1919), 161-271; **48** (1920), 33-94; 208-314.

[F2] P. Fatou, *Sur l'iteration des fonctions transcendantes entières*, Acta Math. **47** (1926), 337-360.

[J] Gaston Julia, *Sur l'iteration des fonctions rationelles*, J. Math. Pures Appl. (7) **4** (1918), 47-245.

[Sk] Wilhelm Schwick, *Repelling periodic points in the Julia set*, Bull. London Math. Soc. **29** (1997), 314-316.

[St] Norbert Steinmetz, *Rational Iteration*, Walter de Gruyter, 1993.

CHAPTER 7

The Prime Number Theorem

The proof of the Prime Number Theorem (PNT) by Jacques Hadamard and (independently) Charles de la Vallée Poussin in 1896 is arguably the high water mark of nineteenth century mathematics. Conjectured on the basis of numerical evidence (independently and in somewhat different forms) by Legendre and Gauss at the end of the eighteenth century, PNT asserts that the number $\pi(x)$ of primes less than or equal to x is asymptotic to $x/\log x$ in the sense that

$$\lim_{x\to\infty}\frac{\pi(x)}{x/\log x}=1.$$

Since the time of Riemann, it has been understood that the distribution of primes is closely connected with the function theoretic properties of the Riemann zeta function $\zeta(s)$, defined initially for $\operatorname{Re} s>1$ by

$$\zeta(s)=\sum_{n=1}^{\infty}\frac{1}{n^s}$$

and extended via analytic continuation to $\mathbb{C}$ as a meromorphic function with a single simple pole at $s=1$. Here the key fact relating the zeta function and PNT is that

$$(*)\qquad \zeta(s)\neq 0 \quad \text{on the line} \quad \operatorname{Re} s=1.$$

The original proofs of PNT involved integration over infinite contours and therefore required, in addition to the nonvanishing of $\zeta(s)$ on $\operatorname{Re} s=1$, certain estimates of $\zeta(s)$ near ∞. Subsequent proofs avoided this difficulty but required instead some version of Wiener's Tauberian theory for Fourier integrals (cf., for instance, the proof using the Wiener-Ikehara theorem given in [**C**]). Thus the deduction of PNT from (*) remained highly nontrivial. In 1980, Donald Newman [**N**] discovered an amazingly simple route to deriving PNT from (*). Newman's innovation, in his own words, was "to return to contour integral methods so as to avoid Fourier analysis, and also to use finite contours so as to avoid estimates at infinity." While Newman applied his method to Dirichlet series, we find it more convenient, following Korevaar [**K**], to use it to prove the following Tauberian theorem for Laplace transforms.

THEOREM. *Let f be a bounded measurable function on $[0,\infty)$. Suppose that the Laplace transform*

$$g(z)=\int_0^{\infty} f(t)e^{-zt}dt,$$

which is defined and analytic on the open half plane $H=\{z:\operatorname{Re} z>0\}$, extends analytically to (an open set containing) $\overline{H}=\{z:\operatorname{Re} z\geq 0\}$. Then the improper

integral $\int_0^\infty f(t)dt = \lim_{T\to\infty} \int_0^T f(t)dt$ *converges and coincides with* $g(0)$, *the value of the analytic extension of* g *at* $z = 0$.

REMARK. This result is not new; in fact, it is a special case of a result of Ingham [**I**], proved by Fourier methods almost half a century earlier. What is of interest here is the simplicity of the proof: by a proper choice of contour and integrand, all previous difficulties are finessed, and one obtains an argument which uses nothing more advanced than the Cauchy integral formula and completely straightforward estimates.

PROOF. Assume that $|f(t)| \leq M$ for all $t \geq 0$. For $T > 0$, the function $g_T(z) = \int_0^T f(t)e^{-zt}dt$ is clearly entire. We claim that

$$\lim_{T\to\infty} g_T(0) = g(0). \tag{7.1}$$

To this end, take $R > 0$ large and $\delta = \delta(R) > 0$ so small that g is analytic on the region $D = \{z : |z| \leq R, \operatorname{Re} z \geq -\delta\}$. Let $\Gamma = \partial D$. Then by Cauchy's Theorem,

$$g(0) - g_T(0) = \frac{1}{2\pi i}\int_\Gamma [g(z) - g_T(z)]e^{zT}\left(1 + \frac{z^2}{R^2}\right)\frac{1}{z}\,dz. \tag{7.2}$$

Let $x = \operatorname{Re} z$. Then for $x > 0$,

$$|g(z) - g_T(z)| = \left|\int_T^\infty f(t)e^{-zt}dt\right| \leq M\int_T^\infty e^{-xt}dt = \frac{Me^{-xT}}{x}, \tag{7.3}$$

while

$$\left|e^{zT}\left(1 + \frac{z^2}{R^2}\right)\frac{1}{z}\right| = e^{xT}\frac{2|x|}{R^2} \quad \text{for} \quad |z| = R. \tag{7.4}$$

Thus, when $z \in \Gamma_+ = \Gamma \cap \{\operatorname{Re} z > 0\}$, the integrand in (7.2) is bounded in absolute value by $2M/R^2$, and hence

$$\left|\frac{1}{2\pi i}\int_{\Gamma_+} [g(z) - g_T(z)]e^{zT}\left(1 + \frac{z^2}{R^2}\right)\frac{1}{z}\,dz\right| \leq \frac{M}{R}. \tag{7.5}$$

On $\Gamma_- = \Gamma \cap \{\operatorname{Re} z < 0\}$, we consider the integrals involving $g(z)$ and $g_T(z)$ separately. Since g_T is entire, we can replace the contour Γ_- by the semicircle $\Gamma'_- = \{z : |z| = R, \operatorname{Re} z < 0\}$. For $x = \operatorname{Re} z < 0$, we have

$$|g_T(z)| = \left|\int_0^T f(t)e^{-zt}dt\right| \leq M\int_{-\infty}^T e^{-xt}dt = \frac{Me^{-xT}}{|x|}; \tag{7.6}$$

so by by (7.4) and (7.6),

$$\left|\frac{1}{2\pi i}\int_{\Gamma'_-} g_T(z)e^{zT}\left(1 + \frac{z^2}{R^2}\right)\frac{1}{z}\,dz\right| \leq \frac{M}{R}. \tag{7.7}$$

Finally, since g is analytic on Γ_-, there exists a constant $K = K(R, \delta)$ such that

$$\left|g(z)\left(1 + \frac{z^2}{R^2}\right)\frac{1}{z}\right| \leq K \quad \text{on} \quad \Gamma_-.$$

Since e^{zT} is bounded on Γ_- and converges uniformly to 0 on compact subsets of $\{\operatorname{Re} z < 0\}$ as $T \to \infty$, it follows easily that

$$\lim_{T\to\infty} \left| \frac{1}{2\pi i} \int_{\Gamma_-} g(z)e^{zT} \left(1 + \frac{z^2}{R^2}\right) \frac{1}{z}\, dz \right| = 0. \tag{7.8}$$

From (7.2), (7.5), (7.7), and (7.8), we have

$$\overline{\lim_{T\to\infty}}\, |g(0) - g_T(0)| \leq \frac{2M}{R}.$$

Since R can be chosen arbitrarily large, this proves (7.1). □

Now let us turn to the actual proof of the Prime Number Theorem, following the concise and elegant development of Zagier [**Z**], which is a model of efficient organization. We begin our discussion with a brief introduction to the Riemann zeta function. Following longstanding tradition, we write the complex variable as $s = \sigma + it$ instead of $z = x + iy$. Define for $\operatorname{Re} s > 1$

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}.$$

Since $|1/n^s| = 1/n^\sigma$, this series converges absolutely for $\sigma > 1$ and uniformly on $\sigma \geq 1 + \varepsilon$ for each $\varepsilon > 0$. Thus, since the functions $1/n^s = e^{-s \log n}$ are all entire, $\zeta(s)$ is analytic for $\operatorname{Re} s > 1$.

LEMMA 7.1. $\zeta(s) - \frac{1}{s-1}$ *extends analytically to* $\operatorname{Re} s > 0$.

PROOF. For $\operatorname{Re} s > 1$,

$$\zeta(s) - \frac{1}{s-1} = \sum_{n=1}^{\infty} \frac{1}{n^s} - \int_1^{\infty} \frac{1}{x^s} dx = \sum_{n=1}^{\infty} \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s}\right) dx.$$

Each summand in the series on the right is evidently an entire function, and the series converges absolutely for $\operatorname{Re} s > 0$ since

$$\left| \int_n^{n+1} \left(\frac{1}{n^s} - \frac{1}{x^s}\right) dx \right| \leq \max_{n\leq x\leq n+1} \left| \frac{1}{n^s} - \frac{1}{x^s} \right| \leq \max_{n\leq u\leq n+1} \left| \frac{s}{u^{s+1}} \right| = \frac{|s|}{n^{\sigma+1}}.$$

Accordingly, convergence is uniform for $\operatorname{Re} s \geq \varepsilon$ for each $\varepsilon > 0$, and so the right hand side is analytic for $\operatorname{Re} s > 0$. □

REMARK. It is not difficult to show that $\zeta(s) - \frac{1}{s-1}$ actually extends to an entire function. However, we do not require this fact.

Henceforth p denotes a prime number, and sums and products over the index p are taken over all primes. The connection between prime numbers and the zeta function is encoded in the next result, known (for real s) already to Euler.

LEMMA 7.2. $\zeta(s) = \prod_p (1 - 1/p^s)^{-1}$ *for* $\operatorname{Re} s > 1$.

PROOF. (Cf. [**A**, p. 213]) Writing p_k for the kth prime, we have

$$\left(1 - \frac{1}{2^s}\right)\left(1 - \frac{1}{3^s}\right) \cdots \left(1 - \frac{1}{p_k^s}\right)\zeta(s) = \sum_{2,3,\ldots,p_k \nmid n} \frac{1}{n^s} \to 1$$

as $k \to \infty$. □

It is easy to see that the Euler product for $\zeta(s)$ converges absolutely for $\operatorname{Re} s > 1$ and uniformly for $\operatorname{Re} s \geq 1+\varepsilon$ for each $\varepsilon > 0$. These facts will be used without further mention below.

Our next result contains the function-theoretic heart of the proof of PNT. Define

$$\Phi(s) = \sum_p \frac{\log p}{p^s}.$$

Since the series converges absolutely for $\operatorname{Re} s > 1$ and uniformly for $\operatorname{Re} s \geq 1+\varepsilon$ for each $\varepsilon > 0$, Φ is analytic in $\operatorname{Re} s > 1$.

LEMMA 7.3. $\Phi(s) - \frac{1}{s-1}$ *extends analytically to* $\operatorname{Re} s \geq 1$, *and* $\zeta(s) \neq 0$ *for* $\operatorname{Re} s = 1$.

PROOF. The proof of Lemma 7.2 shows that $\zeta(s) \neq 0$ for $\operatorname{Re} s > 1$. A simple calculation based on the product representation then yields

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_p \frac{\log p}{p^s - 1} = \Phi(s) + \sum_p \frac{\log p}{p^s(p^s-1)}. \tag{7.9}$$

The last term on the right converges and defines an analytic function for $\operatorname{Re} s > 1/2$, so it follows from Lemma 7.1 that $\Phi(s)$ extends to a meromorphic function on $\operatorname{Re} s > 1/2$ with poles only at $s = 1$ and at the zeros of $\zeta(s)$ and that $\Phi(s) - \frac{1}{s-1}$ is analytic at $s = 1$. Thus, it remains only to show that $\zeta(s)$ does not vanish for $\operatorname{Re} s = 1$.

To this end, recall that if a meromorphic function f vanishes to (exact) order k at s_0, then

$$\lim_{s \to s_0} (s - s_0)\frac{f'(s)}{f(s)} = \operatorname{Res}\left(\frac{f'}{f}, s_0\right) = k \tag{7.10}$$

and, similarly, that if f has a pole of order k at s_0,

$$\lim_{s \to s_0} (s - s_0)\frac{f'(s)}{f(s)} = \operatorname{Res}\left(\frac{f'}{f}, s_0\right) = -k. \tag{7.11}$$

Suppose now that $\zeta(s)$ has a zero of order $\mu \geq 0$ at $s = 1 + i\alpha$ ($\alpha \neq 0$, $\alpha \in \mathbb{R}$); since $\zeta(s)$ is real for real s, it follows that $\zeta(s)$ has a zero of the same multiplicity at $1 - i\alpha$. Denoting the multiplicity of the zeros (if any) at $s = 1 \pm 2i\alpha$ by $\nu \geq 0$ and applying (7.10) and (7.11) to the function $\Phi(s)$, which differs from $-\zeta'(s)/\zeta(s)$ by a function analytic on $\operatorname{Re} s > 1/2$, we obtain

$$\begin{gathered} \lim_{\varepsilon \to 0+} \varepsilon\Phi(1+\varepsilon) = 1 \qquad \text{and} \\ \lim_{\varepsilon \to 0+} \varepsilon\Phi(1+\varepsilon \pm i\alpha) = -\mu \qquad \lim_{\varepsilon \to 0+} \varepsilon\Phi(1+\varepsilon \pm 2i\alpha) = -\nu. \end{gathered} \tag{7.12}$$

But for $\varepsilon > 0$,

$$\sum_{k=-2}^{2} \binom{4}{2+k} \Phi(1+\varepsilon+ik\alpha) = \sum_p \frac{\log p}{p^{1+\varepsilon}} \left(p^{i\alpha/2} + p^{-i\alpha/2}\right)^4 \geq 0, \tag{7.13}$$

since the quantity in parentheses on the right is real. Multiplying (7.13) by ε and using (7.12) to calculate the limit of the left hand side as $\varepsilon \to 0+$, we obtain $-2\nu - 8\mu + 6 \geq 0$. Thus $\mu = 0$, i.e., $\zeta(1 + i\alpha) \neq 0$. This concludes the proof of Lemma 7.3. □

We have completed the preparations for proving PNT. The rest of the proof focuses on the function

$$\theta(x) = \sum_{p \le x} \log p.$$

We shall show that $\theta(x) \sim x$, i.e., $\lim_{x\to\infty} \frac{\theta(x)}{x} = 1$. This easily implies PNT since

$$\theta(x) = \sum_{p\le x} \log p \le \sum_{p \le x} \log x = \pi(x)\log x,$$

while for any $\varepsilon > 0$,

$$\theta(x) \ge \sum_{x^{1-\varepsilon} \le p \le x} \log p \ge \sum_{x^{1-\varepsilon} \le p \le x} (1-\varepsilon)\log x = (1-\varepsilon)\log x[\pi(x) + O(x^{1-\varepsilon})].$$

First, following Chebyshev, we prove

LEMMA 7.4. $\theta(x) = O(x)$.

PROOF. For n a positive integer, we have

$$2^{2n} = (1+1)^{2n} = \sum_{k=0}^{2n} \binom{2n}{k} \ge \binom{2n}{n} \ge \prod_{n<p\le 2n} p = e^{\theta(2n)-\theta(n)},$$

so that $\theta(2n) - \theta(n) \le 2n\log 2$. It follows that

$$\begin{aligned}\theta(x) - \theta(x/2) = \theta(x) - \theta([x/2]) &\le \log x + \theta(2[x/2]) - \theta([x/2]) \\ &\le \log x + 2[x/2]\log 2 \le (1+\log 2)x.\end{aligned}$$

Summing successively over $x, x/2, \dots, x/2^r$, where $2^r > x$, we obtain $\theta(x) \le 2(1+\log 2)x$. □

LEMMA 7.5. *The integral* $\int_1^\infty [\theta(x) - x]/x^2\, dx$ *converges.*

PROOF. This follows directly from the Tauberian theorem of Section 1 applied to the function $f(t) = \theta(e^t)e^{-t} - 1$, which by Lemma 7.4 is bounded. Indeed, using Lemma 7.4 again, we have for $\operatorname{Re} s > 1$,

$$\Phi(s) = \sum_p \frac{\log p}{p^s} = \int_1^\infty \frac{d\theta(x)}{x^s} = s\int_1^\infty \frac{\theta(x)}{x^{s+1}}dx = s\int_0^\infty e^{-st}\theta(e^t)dt,$$

so that

$$\begin{aligned} g(s) = \int_0^\infty f(t)e^{-st}dt = \int_0^\infty [\theta(e^t)e^{-t} - 1]e^{-st}dt &= \frac{\Phi(s+1)}{s+1} - \frac{1}{s} \\ &= \frac{1}{s+1}\left[\Phi(s+1) - \frac{1}{s} - 1\right],\end{aligned}$$

which extends analytically to $\operatorname{Re} s \ge 0$ by Lemma 7.3. Thus

$$\int_1^\infty \frac{\theta(x) - x}{x^2}dx = \int_0^\infty [\theta(e^t)e^{-t} - 1]dt = \int_0^\infty f(t)dt,$$

which converges. □

To complete the proof of PNT, let us show how Lemma 7.5 implies that $\theta(x) \sim x$. Assume that for some $\lambda > 1$, there exist arbitrarily large x with $\theta(x) \geq \lambda x$. Then, since θ is nondecreasing, for each such x,

$$\int_x^{\lambda x} \frac{\theta(t) - t}{t^2}\, dt \geq \int_x^{\lambda x} \frac{\lambda x - t}{t^2} dt = \int_1^{\lambda} \frac{\lambda - t}{t^2} dt > 0,$$

which implies the divergence of $\int_1^\infty [\theta(t) - t]/t^2 dt$, contrary to Lemma 7.5. Similarly, if $\theta(x) \leq \lambda x$ for some $\lambda < 1$ and arbitrarily large x, we would have

$$\int_{\lambda x}^{x} \frac{\theta(t) - t}{t^2}\, dt \leq \int_{\lambda x}^{x} \frac{\lambda x - t}{t^2} dt = \int_{\lambda}^{1} \frac{\lambda - t}{t^2} dt < 0,$$

which would again contradict the convergence of $\int_1^\infty [\theta(t) - t]/t^2 dt$. Thus

$$\lim_{x \to \infty} \theta(x)/x = 1,$$

and the proof is done.

Bibliography

[A] Lars V. Ahlfors, *Complex Analysis*, third edition, McGraw-Hill, 1979.

[C] K. Chandrasekharan, *Introduction to Analytic Number Theory*, Springer-Verlag, 1968.

[I] A.E. Ingham, *On Wiener's method in Tauberian theorems*, Proc. London Math. Soc. (2) **38** (1935), 458-480.

[K] J. Korevaar, *On Newman's quick way to the prime number theorem*, Math. Intelligencer **4** (3) (1982), 108-115.

[N] D.J. Newman, *Simple analytic proof of the prime number theorem*, Amer. Math. Monthly **87** (1980), 693-696.

[Z] D. Zagier, *Newman's short proof of the prime number theorem*, Amer. Math. Monthly **104** (1997), 705-708.

Coda: Transonic Airfoils and SLE

We close by describing two rather unusual applications of complex variables. The details are beyond the scope of this book, but the ideas involved definitely deserve mention.

The first area of application is fluid dynamics. It was observed already in the nineteenth century that the equations describing the incompressibility and irrotationality of fluids are just the Cauchy-Riemann equations for the velocity components in two-dimensional flow. Since low velocity flow is nearly incompressible, this made it possible to use analytic functions (more specifically, the theory of conformal mapping) to describe such flows around airfoils and to determine lift and drag. However, for high speed flows, which are compressible, this approach is not available.

In high speed flows over airfoils, the flow becomes supersonic over parts of the airfoil. This leads to the formation of shock waves, an undesirable effect since shocks increase drag. Although Cathleen Morawetz proved mathematically that, in general, shock waves occur in partially supersonic flows [**M1**], [**M2**], this did not rule out the existence of special airfoils for which shockless flows are possible. In fact, Paul Garabedian and his student David Korn developed a hodograph method based on complex characteristics that enabled them to calculate supercritical wing sections free of shocks at a specified speed and angle of attack [**K**], [**GK1**]. However, the extensive trial and error involved in the selection of parameters defining the flow rendered this method impractical. After the preliminary results of [**BGK**], a completely satisfactory solution of the problem was obtained by Garabedian and Korn in [**GK2**]. They solve the partial differential equations of two-dimensional inviscid gas dynamics by analytic continuation into the domain of two independent complex characteristic coordinates. After mapping the domain of integration conformally onto the unit disk in the plane of one of these coordinates, they formulate a boundary value problem on that disk for the stream function which is well-posed even in the case of transonic flow. This enables them to give a procedure for calculating an airfoil on which the speed is prescribed as a function of arclength, leading to an exact solution of the problem in the case of subsonic flow and, in the transonic case, generally to a shockless flow which assumes the assigned subsonic values of the speed and approximates the given supersonic values. Truly a tour de force of applied complex analysis.

The second area of application is statistical mechanics, and the mathematics has its origin in Charles Loewner's study of univalent (i.e., one-to-one) analytic functions defined on the unit disk. Based on certain known extremal properties of

the function

$$k(z) = \frac{z}{(1-z)^2} = z + \sum_{n=2}^{\infty} nz^n,$$

Bieberbach conjectured that for any univalent analytic function on the unit disk satisfying the normalization

$$f(z) = z + \sum_{n=2}^{\infty} a_n z^n,$$

the coefficient inequality $|a_n| \leq n$ holds, with equality only for $k(z)$ and its rotates $k(\alpha z)/\alpha$, $|\alpha| = 1$. For $n = 2$, this can be demonstrated easily, but for $n \geq 3$ it remained a challenge.

Loewner was able to prove that $|a_3| \leq 3$ by embedding the function f into a one-parameter family of mappings, constructed as follows. Suppose f maps the unit disk onto the exterior of a curve connecting some point p to ∞. Moving the point p along the curve gives a one-parameter family of exterior domains; denote by $f(z;p)$ the (normalized) analytic function mapping the open unit disk onto the exterior of the curve. Loewner **[Lo]** derived a differential equation for f as a function of p and used it successfully to estimate a_3. Loewner's method found significant applications to several other problems in the theory of univalent functions [**D**, pp. 95-117], but efforts to apply it to higher coefficients met with little success; and for the next 60 years, attention was focused on a variety of other approaches to the problem. However, when the Bieberbach Conjecture was finally proved (by Louis de Branges [**Br**]), it was via Loewner's approach; cf. [**FP**].

More recently, Oded Schramm [**S**] discovered a conformally invariant stochastic process, obtained by solving Loewner's equation with Brownian motion as input, which describes scaling limits in statistical mechanics. SLE, the stochastic Loewner evolution (or Schramm-Loewner evolution), was used subsequently to solve many two-dimensional problems in statistical mechanics. To cite but a single example, Lawler, Schramm and Werner [**LSW**] used it to prove Mandelbrot's conjecture that the dimension of the planar Brownian frontier (i.e., the boundary of the infinite connected component of the complement of a planar Brownian path) is 4/3. SLE has led to a major leap in our understanding of the random fractal geometry of such two-dimensional systems as critical percolation and critical Ising models [**Sm1**], [**Sm2**]. It also has close connections with two-dimensional conformal field theory, two-dimensional quantum gravity, and random matrix theory. Surely this work, which figures prominently in two recent Fields Medal citations,[1] is a most striking example of an idea which, originating in the purest mathematics, has turned out to be instrumental in theoretical physics.

[1] To Wendelin Werner (2006) "For his contributions to the development of stochastic Loewner evolution, the geometry of two-dimensional Brownian motion, and conformal field theory" and to Stanislav Smirnov (2010) "For the proof of conformal invariance of percolation and the planar Ising model in statistical physics." Moreover, according to the obituary for Oded Schramm published in the *New York Times* on September 10, 2008, "If Dr. Schramm had been born three weeks and a day later, he would almost certainly have been one of the winners of the Fields Medal ... in 2002."

Bibliography

[BGK] F. Bauer, P. Garabedian and D. Korn, *A Theory of Supercritical Wing Sections*, Lecture Notes in Economics and Mathematical Systems **66**, Springer-Verlag, 1972.

[Br] Louis de Branges, *A proof of the Bieberbach Conjecture*, Acta Math. **154** (1985), 137-152.

[D] Peter L. Duren, *Univalent Functions*, Springer-Verlag, 1983.

[FP] Carl H. FitzGerald and Ch. Pommerenke, *The de Branges theorem on univalent functions*, Trans. Amer. Math. Soc. **290** (1985), 683-690.

[GK1] Paul Garabedian and D.G. Korn, *Numerical design of transonic airfoils*, Numerical Solution of Partial Differential Equations – 2, Academic Press, 1971, pp. 253-271.

[GK2] Paul Garabedian and D.G. Korn, *A systematic method for computer design in supercritical airfoils in cascade*, Comm. Pure Appl. Math. **29** (1976), 369-382.

[K] D.G. Korn, *Computation of shock-free transonic flows for airfoil design*, AEC Research and Development Report NYO-1480-125, Courant Institute of Mathematical Sciences, NYU, 1969.

[La] Gregory F. Lawler, *Conformally Invariant Processes in the Plane*, Amer. Math. Soc., 2005.

[LSW] Gregory F. Lawler, Oded Schramm and Wendelin Werner, *The dimension of the planar Brownian frontier is 4/3*, Math. Res. Lett. **8** (2001), 401-411.

[Lo] Karl Löwner, *Untersuchungen über schlichte konforme Abbildungen des Einheitskreises. I.*, Math. Ann. **89** (1923), 103-121.

[M1] Cathleen S. Morawetz, *On the non-existence of continuous transonic flows past profiles*, Comm. Pure Appl. Math. **9** (1956), 45-68; II, ibid. **10** (1957), 107-131; III, ibid. **11** (1958), 129-144.

[M2] Cathleen S. Morawetz, *Non-existence of transonic flows past a profile*, Comm. Pure Appl. Math. **17** (1964), 357-367.

[S] Oded Schramm, *Scaling limits of loop-erased random walks and uniform spanning trees*, Israel J. Math. **118** (2000), 221-288.

[Sm1] Stanislav Smirnov, *Critical percolation in the plane: conformal invariance, Cardy's formula, scaling limits*, C.R. Acad. Sci. Paris Sér. I Math. **323** (2001), 239-244.

[Sm2] Stanislav Smirnov, *Towards conformal invariance of 2D lattice models*, International Congress of Mathematicians, Vol. II, Eur. Math. Soc., 2006, pp. 1421-1451.

[W] Wendelin Werner, *Random planar curves and Schramm-Loewner evolutions*, Lectures on Probability Theory and Statistics, Lecture Notes in Math. **1840**, Springer, 2004, pp. 107-195.

APPENDIX A

Liouville's Theorem in Banach Spaces

The classical theorem of Liouville asserts that a bounded entire function is constant. There is a corresponding theorem for analytic functions taking values in some complex Banach space X. Recall that a function f defined on a domain $D \subset \mathbb{C}$ and taking values in X is said to be (strongly) analytic on D if

$$f'(z) = \lim_{h \to 0} \frac{f(z+h) - f(z)}{h}$$

exists (in the norm topology) for each $z \in D$. If f is an X-valued function analytic in D and $x^* \in X^*$ is a continuous linear functional defined on X, then it is evident that $x^*(f(z))$ is a complex-valued analytic function of z on D having derivative $x^*(f'(z))$. In particular, if the X-valued function f is entire (i.e., analytic for all $z \in \mathbb{C}$), then $x^*(f(z))$ is an entire function in the classical sense.

EXTENDED LIOUVILLE THEOREM. *Let $F : \mathbb{C} \to X$ be an entire function such that $\|F(z)\|_X \leq M$ for all $z \in \mathbb{C}$. Then there exists $x_0 \in X$ such that $F(z) = x_0$ for all $z \in \mathbb{C}$, i.e., F is constant.*

PROOF. Otherwise, there would exist $z_1, z_2 \in \mathbb{C}$ such that $F(z_1) \neq F(z_2)$, and thus by the Hahn-Banach Theorem, $x^* \in X^*$ such that $x^*(F(z_1)) \neq x^*(F(z_2))$. But for $z \in \mathbb{C}$,

$$|x^*(F(z))| \leq \|x^*\|_{X^*} \|F(z)\|_X \leq M \|x^*\|_{X^*}.$$

Thus $x^*(F(z))$ is a bounded entire function in the classical sense and hence a constant by Liouville's Theorem. This contradicts $x^*(F(z_1)) \neq x^*(F(z_2))$. $\square$

APPENDIX B

The Borel-Carathéodory Inequality

Let F be a function analytic on the closed disc $D = \{z : |z| \leq R\}$. A natural measure of growth of F on D is given by the maximum modulus function

$$M(r) = M(r, F) = \max_{|z| \leq r} |F(z)| = \max_{|z|=r} |F(z)|$$

for $0 \leq r \leq R$. Setting $U(z) = \operatorname{Re} F(z)$ and

$$A(r) = A(r, F) = \max_{|z|=r} U(z),$$

we have the following remarkable inequality, which bounds $M(r)$ in terms of $A(R)$ and $|F(0)|$.

BOREL-CARATHÉODORY INEQUALITY. *Let $0 \leq r < R$. Then*

$$M(r) \leq \frac{2r}{R-r} A(R) + \frac{R+r}{R-r} |F(0)|. \tag{B.1}$$

PROOF. If $F(z) = \sum_{n=0}^{\infty} a_n z^n$, where $a_n = \alpha_n + i\beta_n$ (α_n, β_n real), we have

$$\begin{aligned} U(Re^{i\theta}) &= \operatorname{Re} \sum_{n=0}^{\infty} (\alpha_n + i\beta_n) R^n (\cos n\theta + i \sin n\theta) \\ &= \sum_{n=0}^{\infty} (\alpha_n \cos n\theta - \beta_n \sin n\theta) R^n, \end{aligned}$$

where the series converges uniformly in θ. For $n \geq 1$, we have

$$\begin{aligned} \pi \alpha_n R^n &= \int_0^{2\pi} U(Re^{i\theta}) \cos n\theta \, d\theta \\ \pi \beta_n R^n &= -\int_0^{2\pi} U(Re^{i\theta}) \sin n\theta \, d\theta, \end{aligned}$$

so that

$$\pi a_n R^n = \int_0^{2\pi} U(Re^{i\theta}) e^{-in\theta} d\theta = \int_0^{2\pi} [U(Re^{i\theta}) - A(R)] e^{-in\theta} d\theta.$$

Thus

$$\pi |a_n| R^n \leq \int_0^{2\pi} |U(Re^{i\theta}) - A(R)| d\theta = \int_0^{2\pi} [A(R) - U(Re^{i\theta})] d\theta = 2\pi [A(R) - \alpha_0],$$

so that

$$|a_n| R^n \leq 2[A(R) + [F(0)|] \tag{B.2}$$

and $|a_n|r^n \leq 2[A(R) + |F(0)|](r/R)^n$ for $n \geq 1$. It follows that

$$\begin{aligned}|F(re^{i\theta}) - F(0)| &\leq \sum_{n=1}^{\infty} |a_n|r^n \\ &\leq 2[A(R) + |F(0)|] \sum_{n=1}^{\infty} (r/R)^n \\ &= \frac{2r}{R-r}A(R) + \frac{2r}{R-r}|F(0)|;\end{aligned}$$

hence

$$|F(re^{i\theta})| \leq \frac{2r}{R-r}A(R) + \frac{R+r}{R-r}|F(0)|,$$

as required. □

An immediate consequence is the following general version of Liouville's Theorem.

LIOUVILLE'S THEOREM. *Let $F(z) = U(z) + iV(z)$ be entire and suppose that there exist positive constants C, K, and α such that $U(z) \leq C|z|^\alpha$ whenever $|z| \geq K$. Then $F(z)$ is a polynomial of degree no greater than α.*

PROOF. The hypothesis implies that for each integer $n > \alpha$,

$$\limsup_{R\to\infty} A(R)/R^n \leq 0;$$

so by (B.2), $a_n = 0$ for $n > \alpha$. □

We also have the following characterization of nonvanishing functions of exponential type.

COROLLARY. *Let f be an entire function such that $f(z) \neq 0$ and*

$$\text{(B.3)} \qquad |f(z)| \leq e^{B|z|+C}, \qquad z \in \mathbb{C},$$

for some $B, C > 0$. Then there exist $\alpha, \beta \in \mathbb{C}$ such that $f(z) = e^{\alpha z+\beta}$. If $f(0) = 1$, we may choose $\beta = 0$.

PROOF. Since $f(z) \neq 0$, $f(z) = e^{g(z)}$ for some entire function g. It then follows from (B.3) that $A(r, g) \leq Br + C$; so again by (B.2), g is a linear function. The final assertion of the Corollary is obvious. □

For a comprehensive survey of results related to the Borel-Carathéodory Inequality, see [**KM**].

Bibliography

[KM] Gershon Kresin and Vladimir G. Maz'ya, *Sharp Real-Part Theorems*, Lecture Notes in Math. **1903**, Springer, 2007.

[Z] Lawrence Zalcman, *Picard's Theorem without tears*, Amer. Math. Monthly **85** (1978), 265-268.

APPENDIX C

Phragmén-Lindelöf Theorems

Theorems of Phragmén-Lindelöf type generalize the maximum principle to the situation in which a function f analytic on an unbounded plane domain D remains bounded on the (finite part of) the boundary ∂D. It turns out that if $f(z)$ does not grow too quickly as $z \to \infty$ in D, one may conclude that $|f(z)|$ satisfies the same bound in D as it does on ∂D. The basic result is the following.

THEOREM C.1. *Let f be analytic in the angular region D_α of opening π/α $(\alpha > 1/2)$ between two rays meeting at the origin and continuous on the closed angle. Suppose that $|f(z)| \le M$ on ∂D_α (the union of the rays) and that for some $\beta < \alpha$,*

$$f(re^{i\theta}) = O(e^{r^\beta}) \quad \textit{uniformly in} \quad \theta \quad \textit{as} \quad r \to \infty. \tag{C.1}$$

Then $|f(z)| \le M$ for all $z \in D_\alpha$.

PROOF. Without loss of generality, we may take

$$D_\alpha = \{z = re^{i\theta} : |\theta| < \pi/2\alpha,\ r > 0\}.$$

Fix $\beta < \gamma < \alpha$ and let

$$F_\varepsilon(z) = \exp(-\varepsilon z^\gamma) f(z)$$

for $\varepsilon > 0$. Then

$$|F_\varepsilon(re^{i\theta})| = \exp(-\varepsilon r^\gamma \cos\gamma\theta)|f(re^{i\theta})|. \tag{C.2}$$

Since $\gamma < \alpha$, $\cos\gamma\theta > 0$ for $\theta = \pm\pi/2\alpha$, so $|F_\varepsilon(z)| \le |f(z)| \le M$ for $z = re^{\pm i\pi/2\alpha}$. Moreover, for $z = Re^{i\theta}$ $(|\theta| < \pi/2\alpha)$, we have by (C.1) and (C.2),

$$\begin{aligned} |F_\varepsilon(Re^{i\theta})| &\le \exp(-\varepsilon R^\gamma \cos\gamma\pi/2\alpha)|f(Re^{i\theta})| \\ &\le A\exp(R^\beta - \varepsilon R^\gamma \cos\gamma\pi/2\alpha), \end{aligned}$$

which tends to 0 as $R \to \infty$, since $\gamma > \beta$. Thus, by the maximum principle, $|F_\varepsilon(z)| \le M$ for $z \in D_{\alpha,R} = \{re^{i\theta} : |\theta| < \pi/2\alpha,\ 0 < r < R\}$ and all large R. Letting $R \to \infty$, we see that $|F_\varepsilon(z)| \le M$ in D_α and hence

$$|f(z)| \le M\exp(\varepsilon|z|^\gamma) \quad \text{in} \quad D_\alpha$$

for each $\varepsilon > 0$. Now make $\varepsilon \to 0$ to obtain $|f(z)| \le M$ throughout D_α, as required. □

REMARKS. 1. The full force of assumption (C.1) has not been used in the proof: it clearly suffices for (C.1) to hold for a sequence of values $r = r_n$ with $r_n \to \infty$.

2. Condition (C.1) can be weakened to the requirement that for each $\delta > 0$,

$$f(re^{i\theta}) = O(e^{\delta r^\alpha})$$

uniformly in θ as $r \to \infty$; cf. [**T**, pp. 178-179].

We also have the following result.

THEOREM C.2. *Let f be a bounded analytic function on the doubly infinite strip S and suppose that $|f(z)| \le M$ for $z \in \partial S$. Then $|f(z)| \le M$ for all $z \in S$.*

PROOF. We may assume that $S = \{z : -1 \le \operatorname{Re} z \le 1\}$, so that $|f(\pm 1 + iy)| \le M$ for $-\infty < y < \infty$. Fix $\varepsilon > 0$ and consider the function

$$F(z) = e^{\varepsilon z^2} f(z).$$

Then

$$|F(x+iy)| = e^{\varepsilon(x^2-y^2)}|f(x+iy)|$$

so that, since f is bounded in S,

$$|F(x \pm iT)| \le e^{\varepsilon(1-T^2)}|f(x \pm iT)| \le M$$

for $-1 \le x \le 1$ and T sufficiently large. Thus $|F(z)| \le M$ on the boundary of the rectangle S_T having vertices $\pm 1 \pm iT$ and hence, by the maximum principle, on S_T. Letting $T \to \infty$, we obtain $|F(z)| \le M$ on S; and making $\varepsilon \to 0$ gives $|f(z)| \le M$ there. $\square$

As a simple consequence of Theorem C.2, we have the following analogue of the Hadamard Three Circle Theorem, sometimes attributed to C. Doetsch.

THREE LINES THEOREM. *Let f be a bounded analytic function on the strip $S = \{z : 0 \le \operatorname{Re} z \le 1\}$ and let*

$$M(x) = \sup_{-\infty<y<\infty} |f(x+iy)|, \quad 0 \le x \le 1. \tag{C.3}$$

Then

$$M(x) \le M(0)^{1-x} M(1)^x. \tag{C.4}$$

PROOF. Set $c = \log M(0)/M(1)$. Then by (C.3),

$$|f(z)e^{cz}| \le M(0) \quad \text{for} \quad \operatorname{Re} z = 0 \text{ or } 1.$$

Applying Theorem C.2 to the function $f(z)e^{cz}$ in S, we have

$$|f(x+iy)|e^{cx} \le M(0), \quad 0 \le x \le 1;$$

and from this and the definition of c, (C.4) follows. $\square$

Bibliography

[T] E.C. Titchmarsh, *The Theory of Functions*, second edition, Oxford University Press, 1939.

APPENDIX D

Normal Families

Compactness is undoubtedly one of the "big ideas" of modern analysis. Its application to the study of collections of analytic functions by Paul Montel in his theory of normal families can be taken to mark the birth of modern function theory. Here we recall the main definitions and then state and prove Zalcman's Lemma, which is used in the proof of the Fatou-Julia-Baker Theorem in Chapter 6. As an indication of the efficiency of this approach, we also give a very short proof of the central result of the theory of normal families, Montel's Theorem.

Let D be a domain in the complex plane $\mathbb{C}$. We shall be concerned with analytic maps (i.e., meromorphic functions)

$$f : (D, |\ \ |_{\mathbb{R}^2}) \to (\hat{\mathbb{C}}, \chi)$$

from D (endowed with the Euclidean metric) to the extended complex plane $\hat{\mathbb{C}}$, endowed with the chordal metric χ, given by

$$\chi(z, z') = \frac{|z - z'|}{\sqrt{1 + |z|^2}\sqrt{1 + |z'|^2}}, \qquad z, z' \in \mathbb{C}$$
$$\chi(z, \infty) = \frac{1}{\sqrt{1 + |z|^2}}.$$

Associated to χ is the *spherical derivative*

$$f^{\#}(z) = \lim_{h \to 0} \frac{\chi(f(z+h), f(z))}{|h|}$$
$$= \frac{|f'(z)|}{1 + |f(z)|^2} \qquad (f(z) \neq \infty).$$

Since $\chi(z, w) = \chi(1/z, 1/w)$, $f^{\#} = (1/f)^{\#}$, which provides a convenient formula for $f^{\#}$ at poles of f.

A family $\mathscr{F}$ of meromorphic functions on D is said to be *normal* on D if each sequence $\{f_n\} \subset \mathscr{F}$ has a subsequence which converges χ-uniformly on compact subsets of D. It is easy to see that in case all functions in $\mathscr{F}$ are holomorphic, this condition is equivalent to the requirement that each sequence $\{f_n\} \subset \mathscr{F}$ have a subsequence which either converges uniformly (with respect to the Euclidean metric) on compacta in D or diverges uniformly to ∞ on compacta in D.

Normality is, quite clearly, a compactness notion: a family $\mathscr{F}$ of meromorphic functions on D is normal if and only if it is precompact in the topology of χ-uniform convergence on compact subsets of D. By the Arzelà-Ascoli Theorem, such precompactness is equivalent to the equicontinuity on compacta of the functions in $\mathscr{F}$. And, since these functions are smooth, continuity should be equivalent to the local boundedness of an appropriate derivative. Such is the content of

MARTY'S THEOREM. *A family $\mathscr{F}$ of functions meromorphic on D is normal on D if and only if for each compact subset $K \subset D$, there exists a constant $M(K)$ such that*

$$f^{\#}(z) \leq M(K) \tag{D.1}$$

for all $z \in K$ and all $f \in \mathscr{F}$.

For a proof, see [**A**, pp. 226-227].

Like so many other necessary and sufficient conditions, Marty's Theorem provides less than complete information, principally because condition (D.1) is generally very difficult to verify in those situations in which it is not already evident that the family $\mathscr{F}$ is normal. Accordingly, there has been a continuing search for other conditions which imply normality.

The following result, which has come to be known as Zalcman's Lemma (henceforth ZL), has proved to be particularly useful in this connection.

LEMMA D.1. *A family $\mathscr{F}$ of functions meromorphic on the unit disc Δ is **not** normal if and only if there exist*

(a) *a number $0 < r < 1$*
(b) *points z_n, $|z_n| < r$*
(c) *functions $f_n \in \mathscr{F}$*
(d) *numbers $\rho_n \to 0+$*

such that

$$f_n(z_n + \rho_n \zeta) \to g(\zeta) \tag{D.2}$$

spherically uniformly on compact subsets of $\mathbb{C}$, where g is a nonconstant meromorphic function on $\mathbb{C}$. The function g may be taken to satisfy the normalization

$$g^{\#}(z) \leq g^{\#}(0) = 1 \qquad z \in \mathbb{C}.$$

PROOF. Suppose $\mathscr{F}$ is not normal on Δ. Then by Marty's Theorem, there exists a number r^*, $0 < r^* < 1$, points z_n^* in $\{z : |z| \leq r^*\}$, and functions $f_n \in \mathscr{F}$ such that $f_n^{\#}(z_n^*) \to \infty$. Fix a number $r^* < r < 1$, and let

$$M_n = \max_{|z| \leq r} \left(1 - \frac{|z|}{r}\right) f_n^{\#}(z) = \left(1 - \frac{|z_n|}{r}\right) f_n^{\#}(z_n). \tag{D.3}$$

The maximum exists since $f_n^{\#}$ is continuous for $|z| \leq r$, and it is clear that $M_n \to \infty$. Setting

$$\rho_n = \frac{1}{M_n}\left(1 - \frac{|z_n|}{r}\right) = \frac{1}{f^{\#}(z_n)}, \tag{D.4}$$

we obtain

$$\frac{\rho_n}{r - |z_n|} = \frac{1}{rM_n} \to 0. \tag{D.5}$$

Thus, the functions

$$g_n(\zeta) = f_n(z_n + \rho_n \zeta)$$

are defined for $|\zeta| < R_n$, where $R_n = (r - |z_n|)/\rho_n \to \infty$ as $n \to \infty$. It follows from (D.4) that

$$g_n^{\#}(0) = \rho_n f_n^{\#}(z_n) = 1.$$

For $|\zeta| \le R < R_n$, $|z_n + \rho_n\zeta| < r$, so that by (D.3) and (D.4),

$$g_n^{\#}(\zeta) = \rho_n f_n^{\#}(z_n + \rho_n\zeta) \le \frac{\rho_n M_n}{1 - \frac{|z_n+\rho_n\zeta|}{r}} \le \frac{r - |z_n|}{r - |z_n| - \rho_n R},$$

which tends to 1 as $n \to \infty$ by (D.5). Thus, by Marty's Theorem, $\{g_n\}$ is a normal family. Taking a subsequence, we may assume that the g_n converge uniformly (in the spherical metric) on compact subsets of $\mathbb{C}$ to a meromorphic function g. Clearly, $g^{\#}(\zeta) = \lim g_n^{\#}(\zeta) \le 1$. Finally, g is nonconstant, since $g^{\#}(0) = \lim g_n^{\#}(0) = 1 \ne 0$. It is now evident that if $\mathscr{F}$ consists of analytic functions, the limit function will be entire.

For the converse, assume (a) - (d) and that $\mathscr{F}$ is normal on Δ. By Marty's Theorem, there exists $M > 0$ such that

$$\max_{|z|\le(1+r)/2} f^{\#}(z) \le M$$

for all $f \in \mathscr{F}$. Suppose (D.2) holds and fix $\zeta \in \mathbb{C}$. For large n, $|z_n + \rho_n\zeta| \le (1+r)/2$, so that $\rho_n f_n^{\#}(z_n + \rho_n\zeta) \le \rho_n M$. Thus, for all $\zeta \in \mathbb{C}$,

$$g^{\#}(\zeta) = \lim \rho_n f_n^{\#}(z_n + \rho_n\zeta) = 0.$$

It follows that g is constant (possibly infinity). □

In case $\mathscr{F}$ fails to be normal at $z_0 \in \Delta$, i.e., if $\mathscr{F}$ is not normal in any neighborhood of z_0, we can choose the sequence $\{z_n\}$ in (b) to converge to z_0. The proof of this is itself an amusing application of ZL.

Indeed, suppose that $\mathscr{F}$ is not normal at z_0. Translating if necessary, we may assume that $z_0 = 0$. Of course, 0 is now no longer the center of the disc on which the functions in $\mathscr{F}$ are defined; however, they are all defined in $\{|z| < \rho\}$ for some $\rho > 0$. Take $k_0 \in \mathbb{N}$ so that $1/\sqrt{k_0} < \rho$. Then by Marty's Theorem, for each $k \ge k_0$, there exists $f_k \in \mathscr{F}$ with $\sup\{f_k^{\#}(z) : z \in \Delta(0, 1/2\sqrt{k})\} > k$. For $k \ge k_0$, set $g_k(z) = f_k(z/\sqrt{k})$. Each g_k is defined on $\Delta = \{|z| < 1\}$ and satisfies

$$g_k^{\#}(z) = (1/\sqrt{k}) f_k^{\#}(z/\sqrt{k})$$

there. Clearly, $\sup\{g_k^{\#}(z) : z \in \Delta(0, 1/2)\} > \sqrt{k}$; so again, by Marty's Theorem, $\{g_k\}$ is not normal on Δ. Applying ZL to $\{g_k\}$, we get $0 < r < 1$, $|z_\ell^*| < r$, $\rho_\ell^* \to 0+$ and g_{k_ℓ} such that

$$g_{k_\ell}(z_\ell^* + \rho_\ell^*\zeta) \to g(\zeta)$$

χ-uniformly on compact subsets of $\mathbb{C}$, where $g^{\#}(\zeta) \le g^{\#}(0) = 1$ for $\zeta \in \mathbb{C}$. But this means that

$$f_{k_\ell}\left(\frac{z_\ell^*}{\sqrt{k_\ell}} + \frac{\rho_\ell^*}{\sqrt{k_\ell}}\zeta\right) \to g(\zeta)$$

χ-uniformly on compact subsets of $\mathbb{C}$. Setting $z_\ell = z_\ell^*/\sqrt{k_\ell}$, $\rho_\ell = \rho_\ell^*/\sqrt{k_\ell}$ completes the proof. □

A central result in the theory of normal families is Montel's Theorem, according to which a family of functions meromorphic on a domain D, all of which fail to take on three fixed (and distinct) values in $\hat{\mathbb{C}}$, is normal on D. It is this theorem that makes available the mechanism of normal families for proving global results in (one-dimensional) complex dynamics. Here is a simple and elementary proof of Montel's Theorem, based on ZL; cf. [**R**, pp. 240-241].

MONTEL'S THEOREM. *The collection $\mathscr{F}$ of all meromorphic functions which omit three fixed values $a, b, c \in \hat{\mathbb{C}}$ on a domain $D \subset \mathbb{C}$ is a normal family on D.*

PROOF. Since normality is a local notion, we may suppose that $D = \Delta$, the unit disc. Composing with a linear fractional transformation, we may also assume that the omitted values are $0, 1, \infty$. Let us denote by $\mathscr{F}_n$ the collection of functions on Δ which omit the values $0, \infty$, and all nth roots of 1, so that $\mathscr{F} = \mathscr{F}_1$. Note that $f \in \mathscr{F}$ implies $\sqrt[n]{f} \in \mathscr{F}_n$, while if $h \in \mathscr{F}_n$, then $h^n \in \mathscr{F}$.

Suppose now that $\mathscr{F}$ is *not* normal. Then none of the families $\mathscr{F}_n$ is normal, so by ZL we have, for each n, a nonconstant entire function g_n obtained as a limit of functions omitting all values in $S_n = \{0, 1, e^{2\pi i k/n} : k = 0, 1, \ldots, n-1\}$. By Hurwitz's Theorem, g_n also omits S_n. Moreover, $g_n^{\#}(z) \leq g_n^{\#}(0) = 1$.

Write, for convenience, $T_n = S_{2^n}$, $G_n = g_{2^n}$, and consider the family $\mathscr{G} = \{G_n\}$ on $\mathbb{C}$. Now $G_n^{\#}(z) \leq 1$ for all $z \in \mathbb{C}$, so by Marty's Theorem, $\mathscr{G}$ is normal on $\mathbb{C}$; hence a subsequence converges, χ-uniformly on compacta, to a limit function G. Since $G_n^{\#}(0) = 1$ for all n, $G^{\#}(0) = 1$, so G is nonconstant. The sets T_n are nested, so that G_m omits values in T_n as soon as $m \geq n$. By Hurwitz's Theorem, G must omit T_n for every n. Since $\cup T_n$ is dense in the unit circle and $G(\mathbb{C})$ is an open connected set, this implies that either $G(\mathbb{C}) \subset \Delta$ or $G(\mathbb{C}) \subset \mathbb{C} \setminus \overline{\Delta}$. In either case, we have a contradiction to Liouville's Theorem. □

Immediate (and easy) corollaries of Montel's Theorem include the theorems of Picard, as well as the existence of a direction of Julia for entire functions [**SZ**, p. 352]. The proof just given, together with the standard deduction of Picard's Great Theorem from Montel's Theorem [**SZ**, p. 351], provides the shortest and simplest route to this pinnacle of complex function theory.

COMMENT. Zalcman's Lemma was first stated and proved in [**Z1**]; for a state-of-the-art version, see [**PZ**, Lemma 2]. Additional applications to a wide variety of topics in analysis are in [**Z2**]; see also [**BBHM**], [**Bg**], and [**Bt**]. A survey of various generalizations of Montel's Theorem is given in [**Z3**].

Bibliography

[A] Lars V. Ahlfors, *Complex Analysis*, third edition, McGraw-Hill, 1979.

[BBHM] D. Bargmann, M. Bonk, A. Hinkkanen, and G.J. Martin, *Families of meromorphic functions avoiding continuous functions*, J. Anal. Math. **79** (1999), 379-387.

[Bg] Walter Bergweiler, *A new proof of the Ahlfors Five Islands Theorem*, J. Anal. Math. **76** (1998), 337-347.

[Bt] François Berteloot, *Méthodes de changement d'échelles en analyse complexe*, Ann. Fac. Sci. Toulouse Math. (6) **15** (2006), 427-483.

[PZ] Xuecheng Pang and Lawrence Zalcman, *Normal families and shared values*, Bull. London Math. Soc. **32** (2000), 325-331.

[R] Antonio Ros, *The Gauss map of minimal surfaces*, Differential Geometry, Valencia 2001, World Scientific, 2002, pp. 235-252.

[SZ] Stanislaw Saks and Antoni Zygmund, *Analytic Functions*, third edition, Elsevier, 1971.

[Z1] Lawrence Zalcman, *A heuristic principle in complex function theory*, Amer. Math. Monthly **82** (1975), 813-818.

[Z2] Lawrence Zalcman, *Normal families: new perspectives*, Bull. Amer. Math. Soc (N.S.) **35** (1998), 215-230.

[Z3] Lawrence Zalcman, *Variations on Montel's Theorem*, Bull. Soc. Sci. Lett. Łódź Sér. Rech. Déform. **59** (2009), 25-36.

美国数学会经典影印系列

1. **Lars V. Ahlfors**, Lectures on Quasiconformal Mappings, Second Edition, 978-7-04-047010-9
2. **Dmitri Burago, Yuri Burago, Sergei Ivanov**, A Course in Metric Geometry, 978-7-04-046908-0
3. **Tobias Holck Colding, William P. Minicozzi II**, A Course in Minimal Surfaces, 978-7-04-046911-0
4. **Javier Duoandikoetxea**, Fourier Analysis, 978-7-04-046901-1
5. **John P. D'Angelo**, An Introduction to Complex Analysis and Geometry, 978-7-04-046998-1
6. **Y. Eliashberg, N. Mishachev**, Introduction to the *h*-Principle, 978-7-04-046902-8
7. **Lawrence C. Evans**, Partial Differential Equations, Second Edition, 978-7-04-046935-6
8. **Robert E. Greene, Steven G. Krantz**,
   Function Theory of One Complex Variable, Third Edition, 978-7-04-046907-3
9. **Thomas A. Ivey, J. M. Landsberg**,
   Cartan for Beginners: Differential Geometry via Moving Frames and
   Exterior Differential Systems, 978-7-04-046917-2
10. **Jens Carsten Jantzen**, Representations of Algebraic Groups, Second Edition, 978-7-04-047008-6
11. **A. A. Kirillov**, Lectures on the Orbit Method, 978-7-04-046910-3
12. **Jean-Marie De Koninck, Armel Mercier**, 1001 Problems in Classical Number Theory, 978-7-04-046999-8
13. **Peter D. Lax, Lawrence Zalcman**, Complex Proofs of Real Theorems, 978-7-04-047000-0
14. **David A. Levin, Yuval Peres, Elizabeth L. Wilmer**, Markov Chains and Mixing Times, 978-7-04-046994-3
15. **Dusa McDuff, Dietmar Salamon**, *J*-holomorphic Curves and Symplectic Topology, 978-7-04-046993-6
16. **John von Neumann**, Invariant Measures, 978-7-04-046997-4
17. **R. Clark Robinson**,
    An Introduction to Dynamical Systems: Continuous and Discrete, Second Edition, 978-7-04-047009-3
18. **Terence Tao**,
    An Epsilon of Room, I: Real Analysis: pages from year three of a mathematical blog, 978-7-04-046900-4
19. **Terence Tao**, An Epsilon of Room, II: pages from year three of a mathematical blog, 978-7-04-046899-1
20. **Terence Tao**, An Introduction to Measure Theory, 978-7-04-046905-9
21. **Terence Tao**, Higher Order Fourier Analysis, 978-7-04-046909-7
22. **Terence Tao**, Poincaré's Legacies, Part I: pages from year two of a mathematical blog, 978-7-04-046995-0
23. **Terence Tao**, Poincaré's Legacies, Part II: pages from year two of a mathematical blog, 978-7-04-046996-7
24. **Cédric Villani**, Topics in Optimal Transportation, 978-7-04-046921-9
25. **R. J. Williams**, Introduction to the Mathematics of Finance, 978-7-04-046912-7

《实定理的复证明》是对 Hadamard 的格言“实域中两个真理之间的最好和最短路程是通过复域”的延伸思考。面向熟悉研究生一年级水平分析学的受众，本书的目的在于解释复变量是如何对分析的一些领域中的许多类重要结果提供了快速而高效的证明，这些领域包括诸如近似理论、算子理论、调和分析和复动力系统。

讨论的论题包括了在直线上的加权近似、Müntz 定理、Toeplitz 算子、位移算子的不变空间上的 Beurling 定理、预测理论、Riesz 的凸性定理、Paley-Wiener 定理、Titchmarsh 卷积定理、Gleason-Kahane-Zelazko 定理，以及 Fatou-Julia-Baker 定理。讨论从对代数基本定理的世界上最短的证明开始，而以对素数定理的 Newman 的几乎无果的证明结束。四个简短的附录提供了超过标准的研究生一年级课程所有必要的复分析的背景。热爱分析和漂亮证明的读者将会快乐地一遍又一遍阅读这本薄薄的书并从中受益。

美国数学会经典影印系列

AMS
AMERICAN MATHEMATICAL SOCIETY

本版只限于中华人民共和国境内发行。本版经由美国数学会授权仅在中华人民共和国境内销售，不得出口。

ISBN 978-7-04-047000-0

9 787040 470000 >

定价 55.00 元